AMERICAN

Journal of Mathematics

EDITED BY

FRANK MORLEY

WITH THE COÖPERATION OF

A. COHEN, CHARLOTTE A. SCOTT

AND OTHER MATHEMATICIANS

PUBLISHED UNDER THE AUSPICES OF THE JOHNS HOPKINS UNIVERSITY

Πραγμάτων ἔλεγχος οὐ βλεπομένων

VOLUME XL, NUMBER 3

BALTIMORE: THE JOHNS HOPKINS PRESS

LEMCKE & BUECHNER, *New York.* E. STEIGER & CO., *New York.* WILLIAM WESLEY & SON, *London.*
G. E. STECHERT & CO., *New York.* ARTHUR F. BIRD, *London.* A. HERMANN, *Paris.*

JULY, 1918

Entered as Second-Class Matter at the Baltimore, Maryland, Postoffice.

*Interpolation Properties of Orthogonal Sets of Solutions of Differential Equations.**

BY O. D. KELLOGG.

As the present paper is of the nature of a continuation of the one published in the last number of this Journal, an extended introduction may be dispensed with. Suffice it to say that we shall here be concerned with the problem of extending to the orthogonal function sets arising from ordinary differential equations of second order, the properties there derived for sets arising from integral equations.

1. *The Differential Equation and Boundary Conditions.*

We shall be concerned with differential equations of the form

$$\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right)+(q(x)+\lambda r(x))y=0, \qquad p>0 \text{ for } 0<x<1, \quad r>0 \text{ for } 0\leqq x\leqq 1, \tag{1}$$

with the general homogeneous self-adjoint boundary conditions

$$ay(1)+by'(1)=cy(0)+dy'(0), \quad a'y(1)+b'y'(1)=c'y(0)+d'y'(0). \tag{2}$$

If $p(x)$ vanishes at one of the end points of the interval $(0, 1)$, but in such a way that the differential equation has a solution which remains finite and different from zero, with a finite derivative in the neighborhood of this point, one of the boundary conditions (2) is to be replaced by the condition: $y(x)$ remains finite and different from 0 in the neighborhood of this point. If $p(x)$ vanishes at both end points, both equations (2) are to be replaced by this demand for both end points. Examples of these cases are: (a) the functions $J_0(a_0x)\sqrt{x}$, $J_0(a_1x)\sqrt{x}$, ..., where a_0, a_1, ... are the successive roots of the Bessel function of zero order $J_0(x)$; (b) the Legendre polynomials on the interval $(-1, 1)$.

---

* Presented to the American Mathematical Society, December 1, 1917.

The existence of solutions of the problem (1) and (2), and their oscillation properties have been extensively studied,* and we shall make use of some of the more common ones. Our interest will center rather in the interpolation properties, and some of the consequences of the property (D),† which it will be our object to establish.

By "harmonics" we shall understand solutions of the differential equation and boundary conditions. The corresponding "frequencies" are the values of λ for which these solutions are possible. There is an infinite sequence of frequencies, without finite limit point, and bounded below. We shall think of them written in ascending order of magnitude, a frequency being written twice if it corresponds to two harmonics. The harmonics are orthogonal; ‡ even two corresponding to the same frequency may be made orthogonal by a proper choice of the integration constants entering them. No harmonic vanishes simultaneously with its derivative. The zeros of any harmonic separate those of any harmonic with the same or lower frequency. In addition to these known facts, we assume explicitly (1) the continuity of the coefficients of the differential equation and of the derivatives involved, (2) that the boundary conditions are such that for any pair of harmonics

$$p(0)\begin{vmatrix}\phi_i(0), \phi_j(0)\\ \phi_i'(0), \phi_j'(0)\end{vmatrix} = p(1)\begin{vmatrix}\phi_i(1), \phi_i(1)\\ \phi_j'(1), \phi_j'(1)\end{vmatrix}, \tag{3}$$

and (3), that the i-th *harmonic has exactly* i *zeros in the interior of* $(0, 1)$ *for all* i. A harmonic is "even" or "odd" according as the number of interior zeros is even or odd.

2. *Types of Boundary Conditions.*

For what follows it will be convenient to separate the boundary conditions into types according to the behavior of the corresponding harmonics at the end points of $(0, 1)$. Consider first the case in which the determinant of coefficients in (2) $ab'-a'b$ vanishes. There will then be a relation of the form $a''\phi_i(0)+b''\phi_i'(0)=0$. From this it follows that the function

$$M(x)=p(x)\,[\phi_i'(x)\phi_j(x)-\phi_i(x)\phi_j'(x)]$$

---

* See Bôcher, "Leçons sur les méthodes de Sturm," Borel monographs, Paris, 1917. Rich indications as to the literature are found in the footnotes.

† AMERICAN JOURNAL OF MATHEMATICS, Vol. XXXVIII, No. 1 (1916), particularly (1), (2), I, II, IV, V, VI and V′. Sturm (*Liouville's Journal*, Vol. I, p. 433), and Liouville (same Journal, Vol. I, p. 269) studied the interpolation problem, and the method of the latter has suggested that used here. But the only existing results appear to be confined to boundary conditions of Class I below, which does not include the periodic and other interesting cases.

‡ Orthogonal here in the sense that $\int_0^1\phi_i(x)\phi_j(x)r(x)\,dx=0,\quad i\neq j$.

vanishes for $x=0$ for every i and j, and in consequence, by (3), also for $x=1$. We are thus led to the first class of boundary conditions, in which is also included the cases in which $p(x)$ vanishes at one or both end points.

CLASS I. $M(0)\equiv M(1)\equiv 0$. One boundary condition involves only one end point, and the other, only the other. Consider first the case in which $p(x)$ vanishes at an end point. Then under our assumption, all harmonics are different from 0 at that end point. Consider an end point at which $p(x)$ does not vanish, say $x=0$. Then, as $M(0)=0$, $\phi_i'(0)\phi_j(0)-\phi_i(0)\phi_j'(0)=0$, and since no harmonic vanishes simultaneously with its derivative, it follows that if any harmonic vanishes for $x=0$, all do. We conclude for Class I: *If any harmonic vanishes at an end point, all vanish there.*

Supposing the determinant $cd'-c'd=0$ leads to the same class. We may therefore assume that both this and $ab'-a'b$ are different from zero, and that $p(0)>0$ and $p(1)>0$. The conditions (2) may now be given the form

$$y(1)=ay(0)+by'(0), \tag{4_1}$$
$$y'(1)=cy(0)+dy'(0), \tag{4_2}$$

or

$$y(0)=\quad dy(1)-by'(1), \tag{5_1}$$
$$y'(0)=-cy(1)+ay'(1), \tag{5_2}$$

where, by (3) and (4), $\Delta=ad-bc=p(0)/p(1)>0$. These conditions characterize the second class of cases.

CLASS II. $M(0)\equiv M(1)\not\equiv 0$. The boundary conditions both involve both end points. $M(0)$ and $M(1)$ may both vanish, but not as a consequence of the boundary conditions. There will be three types to consider in this class.

(a) $b=0$. Then $ad>0$, and (4_1) shows that if $\phi_i(x)$ vanishes at either end point, it does at the other. The sign of d now becomes important.

(a_1) $b=0$, $a>0$, $d>0$. If $\phi_i(x)$ vanishes at the end points, (4_2) shows that $\phi_i'(0)$ and $\phi_i'(1)$ have the same signs, so that $\phi_i(x)$ must have an odd number of interior zeros. But (4_1) shows the converse. Hence, for this case: *All odd harmonics vanish, and all even harmonics are different from zero at both end points.* An example of this case is $y''+\lambda y=0$, $y(1)=y(0)$, $y'(1)=y'(0)$, with the harmonics 1, $\sin 2\pi x$, $\cos 2\pi x$, $\sin 4\pi x$, $\cos 4\pi x$, ...

(a_2) $b=0$, $a<0$, $d<0$. By similar reasoning, we conclude for this case: *All even harmonics vanish, and all odd harmonics are different from zero at both end points.* An example is $y''+\lambda y=0$, $y(1)=-y(0)$, $y'(1)=-y'(0)$, with the harmonics $\sin \pi x$, $\cos \pi x$, $\sin 3\pi x$, $\cos 3\pi x$, ...

(b) $b>0$. Since a harmonic will not vanish simultaneously with its derivative, (4_1) shows that no harmonic vanishes at both end points. If

$\phi_i(0)=0$, (4_1) shows that $\phi_i(1)$ and $\phi_i'(0)$ have the same signs, and the harmonic must be even. The same conclusion follows from (5_1) when $\phi_i(1)=0$. Hence in this case: *No odd harmonic vanishes at either end point, and no harmonic vanishes at both.* An example is $y''+\lambda^2 y=0$, $\pi y(1)=2y'(0)$, $2y'(1)=-\pi y(0)$. The harmonics are, for $i\geqq 2$, $\phi_i(x)=\sin\lambda_i(x-1)+\frac{2}{\pi}\cos\lambda_i x$, where λ_i is the i-th positive root of $4\pi\lambda=(\pi^2+4\lambda^2)\sin\lambda$; while $\lambda_0=\lambda_1=\pi/2$. For these frequencies we have the solutions $A\cos\frac{\pi}{2}x+B\sin\frac{\pi}{2}x$. It will be seen if A and B are so chosen as to give an odd harmonic, this harmonic will be different from zero at both end points. All later harmonics are different from zero at both ends.

(c) $b<0$. By similar reasoning we conclude: *No even harmonic vanishes at either end point, and no harmonic does at both.*

3. *Lemmas.*

We shall need the following:

1. *If two harmonics have the same frequency, the one with the greater number of interior zeros is different from zero at both end points.* For the zeros of the two separate each other.

2. Let $y(x)=c_0\phi_0(x)+c_1\phi_1(x)+\ldots\ldots+c_n\phi_n(x)$ be a linear combination of harmonics, in which at least one of the constants c_i before the last is not 0. *If, for $c_n=0$, $y(x)$ has p interior roots, then for some $c_n\neq 0$, it will have at least p interior roots. If for $c_n=0$ it has an infinite number of interior roots, then for some $c_n\neq 0$, it will have at least N interior roots, N being any positive integer.* The proof is easily supplied by considering the limit of $y(x)$ as $c_n \doteq 0$, the approach of $y(x)$ to its limit being uniform. We shall reserve the notation $y(x)$ for functions of the above form.

3. *Given a function $y(x)$, $c_n\neq 0$, c_n remaining fixed, we can choose $c_0, c_1, \ldots, c_{n-1}$ so small, not necessarily zero, that $y(x)$ has the same number of interior zeros as $\phi_n(x)$ (i. e., exactly n), provided $\phi_n(x)$ does not vanish at an end point at which an earlier harmonic is different from zero. And in any case, $y(x)$ has not more than one additional interior zero corresponding to a terminal zero of $\phi_n(x)$.* The proof depends upon the facts that $\phi_n(x)$ does not vanish simultaneously with its derivative, and that all the harmonics are finite, with finite derivatives.

4. *Establishment of the Property* (D).

The results and treatment differ for the different types of boundary conditions enumerated. But the general argument is this. With a solution $\psi(x)$ of the differential equation, corresponding to boundary conditions to be assigned later, and not vanishing in the interior of $(0, 1)$, we form the differential operators to be applied to $y(x)$:

$$Ny(x) = p(x)\,[y'(x)\psi(x) - y(x)\psi'(x)], \text{ and } Ly(x) = \frac{d}{dx}Ny(x). \qquad (6)$$

It will first be shown that repeated application of the operator L reduces relatively the earlier coefficients of $y(x)$ to any desired degree, and thus, by the third lemma, leads to a function with no more than n interior roots. It will then be shown that the application of the operator L does not diminish the number of interior roots. The result, with the modifications noted, will be that no $y(x)$ has more than n interior roots, and accordingly, that the determinant of the equations $y(x_0)=0$, $y(x_1)=0, \ldots, y(x_n)=0$ does not vanish for any set of arguments x_i, no two of which are equal. The property (D) will follow (see this Journal, Vol. XL, p. 146, footnote).

Note first, however, that $D_0(x_0)>0$, since the first harmonic has no interior zero. Second, that in case $\lambda_1=\lambda_0$, $D_1(x_0, x_1)>0$. For the two harmonics, belonging to the same frequency, will have a non-vanishing Wronskian, from which fact the desired result may be inferred. We may therefore suppose in the following, $n \geqq 1$, or, in case $\lambda_1=\lambda_0$, $n\geqq 2$.

The choice of $\psi(x)$ in the operators (6) may vary with the function $y(x)$ to which it is to be applied, and the corresponding parameter value λ^* for which $\psi(x)$ is a solution of the differential equation (1) will vary with it. We shall assume, however, that there are two constants, $\varkappa$ and ρ, such that always

$$\varkappa \leqq \lambda^* \leqq \rho \text{ where } \rho < \frac{\lambda_0+\lambda_1}{2}, \text{ or if } \lambda_1=\lambda_0,\ \rho<\frac{\lambda_0+\lambda_2}{2}. \qquad (7)$$

The assumption will be considered below as use is made of it. Using the differential equation, it will be found that

$$Ly(x) = p(x)\psi(x)r(x)\,[c_0(\lambda^*-\lambda_0)\phi_0(x) + c_1(\lambda^*-\lambda_1)\phi_1(x) + \ldots + c_n(\lambda^*-\lambda_n)\phi_n(x)].$$

Dividing, on the assumption $c_n \neq 0$, by $p(x)\psi(x)r(x)(\lambda^*-\lambda_n)$ we arrive at a new function

$$y(x) = c_0\left(\frac{\lambda^*-\lambda_0}{\lambda^*-\lambda_n}\right)\phi_0(x) + c_1\left(\frac{\lambda^*-\lambda_1}{\lambda^*-\lambda_n}\right)\phi_1(x) + \ldots + c_{n-1}\left(\frac{\lambda^*-\lambda_{n-1}}{\lambda^*-\lambda_n}\right)\phi_{n-1}(x) + c_n\phi_n(x).$$

Under the assumption (7) the quantities $\left(\frac{\lambda^*-\lambda_i}{\lambda^*-\lambda_n}\right)$ are all less in absolute value than certain quantities independent of λ^* and less than 1, an exception arising only in the case of the last of them when $\lambda_n=\lambda_{n-1}$. Hence, by repeating the application of L and the division, we may reduce all the coefficients of $y(x)$ except the last (or, in case $\lambda_n=\lambda_{n-1}$, the last two) to be less than any assigned positive quantity. This completes our first step.

The next step is to study how this process affects the number of interior zeros in the various cases enumerated.

CLASS I. Choose $\psi(x)=\phi_0(x)$. Then $\lambda^*=\lambda_0$, and (7) is satisfied. If $y(x)$ has p interior zeros, $Ny(x)$ will be seen to vanish between each pair, and thus have at least $p-1$ interior zeros. But since in this case $M(x)$ vanishes at both end points for all pairs of harmonics, and as $y(x)$ is linear in the harmonics, it follows that $Ny(x)$ also vanishes at both end points, and hence $Ly(x)$ has at least p interior zeros. Thus no interior zeros are lost by the repeated application of L. Lemma 2 permits us to assume $c_n\neq 0$, and then Lemma 3 shows that no $y(x)$ has more than n interior zeros, since by § 2, no harmonics vanish at an end point under the present boundary conditions unless all do. And a case of double frequency does not arise.* *Hence, in problems of Class I, the property* (*D*) *holds.*

CLASS II, (a_1). $b=0$, $a>0$, and $y(1)=ay(0)$, for $y(x)$ is a linear homogeneous combination of harmonics, and hence satisfies the same boundary conditions. Hence $y(0)$ and $y(1)$, if not both zero, have the same signs. We again use $\psi(x)=\phi_0(x)$, and assume that the last harmonic in $y(x)$, $\phi_n(x)$ is even. If $y(x)$ has an odd number, $2p-1$, of interior zeros, it also has two at the end points, and $Ny(x)$ has therefore at least $2p$ interior zeros, and $Ly(x)$ at least $2p-1$. Thus, if the application of L to $y(x)$ changes this function from one with an odd number of interior zeros to a new one with an odd number, none are lost. If it changes $y(x)$ to a function with an even number of interior zeros, at least one is *gained.* If $y(x)$ has an even number, $2p$, of interior zeros, $Ny(x)$ has at least $2p-1$, but (3) shows that it then either has an additional interior zero, or else two at the end points, and in either case $Ly(x)$ has at least $2p-1$ interior zeros. If $Ly(x)$ has an even number of interior zeros, it must be at least $2p$, and if odd, at least $2p-1$. But ultimately, by Lemma 3, the repeated application of L leads to a function with an even number of interior zeros, since we have supposed this to be the case with

* For two harmonics of the same frequency, $M(x)=\text{const}$. This constant is not zero if the harmonics are independent. But it must be 0 in the problems of Class I.

$\phi_n(x)$. Hence no zeros have been lost, and we conclude: *The harmonics of Case II* (a_1) *have the property* (D) *for every even* n.

The conclusion is still justified in case $\lambda_{n-1}=\lambda_n$. Here we must study $c_{n-1}\phi_{n-1}(x)+c_n\phi_n(x)$. If this has an even number of interior zeros, their number can not exceed n, since they are separated by the zeros of $\phi_{n-1}(x)$, and $c_{n-1}\phi_{n-1}(x)+c_n\phi_n(x)$ will, by the boundary conditions, be different from 0 at the end points. The conclusion thus subsists in this case. If $c_{n-1}\phi_{n-1}(x)+c_n\phi_n(x)$ has an odd number of interior zeros, the repeated application of L to $y(x)$ may have reduced the number of interior zeros of $y(x)$ by one, but by no more, and $c_{n-1}\phi_{n-1}(x)+c_n\phi_n(x)$ has $n-1$ interior zeros, and none at the end points, since by § 2, $\phi_{n-1}(x)$ vanishes at both end points, and $\phi_n(x)$ at neither, and $c_n\neq 0$. The application of the lemma 3 then shows that $y(x)$ can not have more than n interior zeros, and the conclusion is established also in the case $\lambda_{n-1}=\lambda_n$.

Class II, (a_2). $b=0$, $a<0$, and $y(1)=ay(0)$, so that if $y(1)$ and $y(0)$ are different from zero, they have opposite signs. We now operate with L in which $\psi(x)$ is chosen subject to the condition that $y'(x)\psi(x)-y(x)\psi'(x)$ vanishes at both end points. By § 2, $\phi_0(x)$, being an even harmonic, vanishes at both end points, and as $\psi(x)$ is different from zero on the interior of $(0,1)$, it follows that $\lambda^*\leqq\lambda_0$. Hence the assumption (7) is justified as far as the upper limit on λ^* is concerned. The existence of the lower limit will be generally established at the end of the paper. We start with a function $y(x)$ ending in an odd harmonic. The operator N does not reduce the number of interior zeros by more than one, while with the present choice of $\psi(x)$, $Ny(x)$ has zeros at the end points. So the operator L does not diminish the number of interior roots. As the odd harmonic $\phi_n(x)$, by § 2, is different from 0 at the end points, Lemma 3 gives the result: *The harmonics of Case II* (a_2) *have the property* (D) *for every odd* n.

Cases II, (b) and (c). Here, again, we choose $\psi(x)$ so as to make $Ny(x)$ vanish at the end points. It is not, in these cases, apparent that the upper limit for λ^* in (7) holds. However, since $\psi(x)$ does not vanish within $(0,1)$, and as $\phi_2(x)$ vanishes twice, it is clear that $\lambda^*<\lambda_2$. As the frequencies have no finite limit point, there must be a positive integer n, and a number ρ, independent of λ^*, such that $\lambda^*\leqq\rho<\dfrac{\lambda_0+\lambda_n}{2}$. The least value of n for which these inequalities are possible,* we will call n'. The type of argument used in the

* This least value of n can often be easily determined in a particular problem, it being helpful to observe that the value of λ for which the differential equation has a solution vanishing at the end points of the interval (0, 1) but not in the interior, is an upper bound for λ^*. In the example in § 9 under Case II (b), (7) is fulfilled without alteration.

previous cases then permits us to conclude: *The harmonics of Case II (b) have the property (D) for every odd* $n \geqq n'$, *and, provided none of the harmonics after* $\phi_{n'-1}(x)$ *vanishes at either end point, for every* $n \geqq n'$. *Also, the harmonics of Case II (c) have the property (D) for every even* $n \geqq n'$, *and, provided none of the harmonics after* $\phi_{n'-1}(x)$ *vanishes at either end point, for every* $n \geqq n'$.

5. *Existence of a Lower Bound for* λ^*.

In the problems of Class II, from (a_2) inclusive, on, it was assumed that a lower bound for the values of λ^* existed. We proceed to show that such is the case. It will be recalled that $\psi(x)$ was to be chosen so that the ratios $\psi'(1)/\psi(1)$ and $\psi'(0)/\psi(0)$ coincided with the corresponding ratios for $y(x)$. Let us call these ratios η and ξ respectively. They are not independent, but are connected by the equation, obtained by dividing (4_2) by (4_1):

$$\eta = (c+d\xi)/(a+b\xi), \quad \text{where } ad-bc>0. \tag{8}$$

We shall now prove the theorem:

There exists a number Λ, *independent of* ξ *and* η, *such that the solution of the differential equation* (1) *which satisfies the boundary conditions*

$$u'(0)=\xi u(0), \quad u'(1)=\eta u(1), \tag{9}$$

subject to the relation (8), *and which does not vanish in the interior of* (0, 1), *belongs to a parameter value* $\lambda^* > \Lambda$.

We assume that $p(0)>0$ and $p(1)>0$. The theorem is not needed in the problems of Class I.

We first prove the lemma: *For fixed* ξ *and decreasing* η, *and for fixed* η *and increasing* ξ, *the value of the parameter* λ *corresponding to* (9) *increases.* Let $u_1(x)$ and $u_2(x)$ be two solutions of the differential equation satisfying the first equation (9). Denoting the corresponding parameter values by λ_1 and λ_2, we find the following identity:

$$p(1)\,[u_2(1)u_1'(1)-u_1(1)u_2'(1)] = (\lambda_2-\lambda_1)\int_0^1 u_2(x)u_1(x)r(x)\,dx.$$

We are, in this paragraph, concerned only with solutions of (1) which do not vanish in the interior of (0, 1), so that it is legitimate to consider $u_1(x)$ and $u_2(x)$ positive for $0<x<1$, so that the integral is positive. Dividing by $u_2(1)u_1(1)$, we have

$$p(1)\,[\eta(\lambda_1)-\eta(\lambda_2)] = \frac{(\lambda_2-\lambda_1)}{u_2(1)u_1(1)}\int_0^1 u_2(x)u_1(x)r(x)\,dx, \tag{10}$$

which shows that decreasing η corresponds to increasing λ. The second part of the lemma is proved similarly.

We next find the relation between ξ and η when these are the values of the ratios $u'(0)/u(0)$ and $u'(1)/u(1)$ corresponding to any *fixed* λ. To do this we take a fundamental solution set for $x=0$, $u_1(x)$ and $u_2(x)$, for which $u_1(0)=1$, $u_1'(0)=0$, $u_2(0)=0$, $u_2'(0)=1$, in terms of which any solution $u(x)$ for the same value of λ may be expressed in the form $u(x)=c_1 u_1(x)+c_2 u_2(x)$. Then $\xi=c_2/c_1$, and $\eta=[c_1u_1'(1)+c_2u_2'(1)]/[c_1u_1(1)+c_2u_2(1)]$. Eliminating the constants, we find the relation

$$\eta=\frac{u_1'(1)+\xi u_2'(1)}{u_1(1)+\xi u_2(1)}, \tag{11}$$

for which we have the easily derived relation

$$u_1'(1)u_2(1)-u_1(1)u_2'(1)=p(0)/p(1)>0.$$

This shows that the hyperbola which is the Cartesian representation of the equation (11) is an always rising one. For various values of λ, this hyperbola varies in position and size, but if for a finite value Λ of λ, its center lies above and to the left of the hyperbola (8), also an always rising hyperbola, the lemma shows that the solution of the differential equation and the boundary conditions (9) subject to (8) will always correspond to a value of $\lambda>\Lambda$, and the theorem will be proved.

That for some finite λ the center of the hyperbola (11) lies above and to the left of the upper left-hand branch of the hyperbola (8), will be evident from the fact which we shall next prove: given any finite positive quantities h and k, a finite λ always exists for which the coordinates of the center of the hyperbola (11) satisfy the inequalities

$$u_1(1)/u_2(1)>h, \quad u_2'(1)/u_2(1)>k. \tag{12}$$

The identity $\dfrac{u_2(1)}{u_1(1)}=\displaystyle\int_0^1\frac{p(0)}{p(x)u_1^2(x)}\,dx$, the derivation of which gives no trouble if the initial values of the fundamental solution set are kept in mind, shows that the ratio $u_1(1)/u_2(1)$ can be made as great as we wish if $u_1(x)$ can be made as great as we please at all interior points. But this fact can be inferred from the identity

$$\frac{d}{dx}\left(\frac{u_1(x,\lambda)}{u_1(x,\lambda_1)}\right)=\frac{\lambda_1-\lambda}{p(x)u_1^2(x,\lambda_1)}\times\int_0^x u_1(x,\lambda)u_1(x,\lambda_1)r(x)\,dx.$$

This shows that $u_1(x,\lambda)/u_1(x,\lambda_1)$ increases with x if $\lambda<\lambda_1$. This ratio

approaches 1 as $x \doteq 0$. Hence always, $u_1(x, \lambda) > u_1(x, \lambda_1)$ for $0 < x < 1$, and using the law of the mean, $\frac{d}{dx} \frac{u_1(x, \lambda)}{u_1(x, \lambda_1)} > \frac{\lambda_1 - \lambda}{\text{max. } p} \int_0^x r(x)\,dx > (\lambda_1 - \lambda) \times$ a positive function of x independent of λ. Hence the ratio, and therefore $u_1(x, \lambda)$ can be made as large as we please, and the first inequality (12) is established.

To prove the second, we start by noticing that for sufficiently small λ, $u_2'(x) > 0$ for $0 < x < 1$. This is seen from the equation $p(x) u_2'(x) = p(0) + \int_0^x (-\lambda r - q) u_2 dx$, in which the integral is positive if $\lambda < -\text{max. } q/\text{min. } r$. Multiplying the differential equation for u_2 by $2p(x) u_2'(x)$ and integrating, we have $[p u_2'(x)]^2 \Big|_0^1 = -2 \int_0^1 p(\lambda r + q) u_2 u_2' dx = -\overline{p(\lambda r + q)}\, u_2^2(1)$, the bar denoting a mean value. Hence $u_2'(1)/u_2(1) > \frac{1}{p(1)} \sqrt{(-\lambda)\ \text{min. } pr - \text{max. } pq}$, which can be made as large as we please by sufficiently diminishing λ. Thus the second inequality (12) is proved.

COLUMBIA, MO., *October* 20, 1917.

Directed Integration.

By H. B. Phillips.

In case of an integral along a curve,

$$\int P dx + Q dy = \text{Lim}\ \Sigma P \Delta x + Q \Delta y,$$

the increments Δx and Δy may be positive or negative according as x and y are increasing or decreasing. In case of a double integral, however,

$$\int\int f(x, y)\, dx dy = \text{Lim}\ \Sigma\Sigma f(x, y)\, \Delta x \Delta y,$$

the element $\Delta x \Delta y$ is usually considered positive, or at least, invariable in sign. This introduces difficulties similar to those which occur when we attempt to banish the minus sign from algebra or analytic geometry. For instance, in a change of variable, it is necessary to assume that the Jacobian has an invariable sign.

Physicists avoid these difficulties by introducing a cosine which is positive or negative as required. Mathematicians accomplish the same result by making a change of variable, thus obtaining an element of integration which need not change sign. I wish to show in this paper how the algebraic sign can be directly attached to the element of integration, multiple integrals being treated in this respect like curvilinear integrals. The equations for change of variable and those connecting line, surface, and volume integrals, present themselves much more naturally in this form. In this discussion I shall not enter into questions of existence and convergence. These matters are treated in practically the same way whether the integral is directed or not. I shall also consider only two and three dimensions, although the extension to higher spaces is immediate.

Directed Regions.—A surface is called one-sided if it is possible to pass from a point on one side of the surface to a point on the other without passing through the surface or across its border. If this is not possible, the surface is called two-sided. A simple one-sided surface can be formed by twisting a strip of paper through 180° and bringing its ends together.

If a surface is two-sided, one side can be considered positive, the other negative. We shall assign a direction or sense to a region on a two-sided surface by assigning a direction around its border. In a right-hand system the positive direction is usually chosen such that an observer on the positive side of the surface finds the region on his left when he moves in the positive direction along the border. It should be noted that this direction does not belong to the border, but to the region of which it is the border. Thus a given direction around a great circle of a sphere is positive for the hemisphere on one side, and negative for that on the other.

Surface Integrals.—Let x and y be one-valued and continuous functions defined at each point of a region Γ on a plane or two-sided surface. Divide Γ into elementary regions, or cells, by two sets of curves

$$x = \text{constant}, \quad y = \text{constant}.$$

Any one of these cells whose boundary is a simple quadrilateral with two pairs of opposite sides belonging to the curves x, $x+\Delta x$ and y, $y+\Delta y$ will be called regular. Irregular cells may be bounded by less than four curves, or by four curves that are not of this simple type.

In the definition of the integral only the regular cells will be used. Hence it is assumed that the irregular cells can be enclosed in a region or set of regions whose total area approaches zero when Δx and Δy approach zero. This is certainly true of the curves ordinarily used in integration. It would not be true if the two systems of curves $x=$const. and $y=$const. were the same.

Choose a direction around one of the quadrilaterals. Then for that quadrilateral we define $\Delta x\Delta y$ as the product obtained by multiplying the increments of x and y which are found by passing around the quadrilateral in the chosen direction so as first to traverse a curve $y=$const., and then a curve $x=$const. The sign of $\Delta x\Delta y$ is fixed for a given quadrilateral and a given direction around it. Thus in the quadrilateral $ABCD$, if AB and DC are portions of curves $y=$const., and AD and BC portions of curves $x=$const., we may take Δx from A to B and Δy from B to C, or we may take Δx from C to D and Δy from D to A. In the second case the signs of Δx and Δy are both changed, and so $\Delta x\Delta y$ is not changed. Similarly, the product $\Delta y\Delta x$ is obtained by traversing the quadrilateral in the same direction, first traversing a curve $x=$const. and then a curve $y=$const. Inspection of a figure will make it clear that one of the increments in $\Delta y\Delta x$ differs in sign from the corresponding increment in $\Delta x\Delta y$. Hence

$$\Delta y\Delta x = -\Delta x\Delta y. \tag{1}$$

Let the same direction be taken around all the quadrilaterals into which Γ is divided. Then, if (x, y) is any point in the quadrilateral to which $\Delta x \Delta y$ belongs, we define the integral of $f(x, y)$,

$$\int\int f(x, y)\, dxdy,$$

in the chosen direction over Γ as the limit (if it exists) approached by the sum

$$\Sigma\Sigma f(x, y)\, \Delta x \Delta y$$

when Δx and Δy approach zero, the summation being for all the regular cells within Γ. Similarly,

$$\int\int f(x, y)\, dydx = \text{Lim } \Sigma\Sigma f(x, y)\, \Delta y \Delta x.$$

Therefore, by equation (1),

$$\int\int f(x, y)\, dydx = -\int\int f(x, y)\, dxdy. \tag{2}$$

It should be noted that the order in which the differentials are written does not indicate an order of integration. In fact, no order of integration is considered. The integrals are multiple, not iterated.

We have assumed that Δx is determined along the curves $y=$const., and Δy along the curves $x=$const. It is a very important fact that one of these increments could be determined along a third set of curves $w=$const. Thus, if we resolve Γ into cells by the curves $x=$const. and $w=$const., and in each quadrilateral determine Δy on $x=$const. as before, but Δx on $w=$const., the value of the integral will not be changed, provided the total area of the irregular cells formed by the new curves has a zero limit. For Γ can be resolved into strips between consecutive curves x, $x+\Delta x$. All the quadrilaterals in a strip have the same Δx. Also Δy is taken in both cases along the curves $x=$const. Hence, in the change assumed, the part of the summation belonging to this strip is affected only through the change in the distribution of the intervals Δy. This does not affect the limit.

Volume Integrals.—Triple integrals are defined in a similar way. Let x, y, z be one-valued and continuous functions defined at each point of a region Γ of space. Divide Γ into cells by means of three sets of surfaces $x=$const., $y=$const., and $z=$const. We shall call the cells regular which are bounded on opposite sides by three pairs of surfaces x and $x+\Delta x$, y and $y+\Delta y$, z and $z+\Delta z$. We assume that when Δx, Δy, and Δz approach zero, the total volume of the irregular cells approaches zero.

Let AB, BC, CD be consecutive edges of a cell, y and z being constant on AB, z and x on BC, x and y on CD. Let the outer surface of the cell be considered positive. The path BCD determines a direction (positive or negative)

about the face of the cell in which B, C, D lie. I call this the direction, or sense, of the cell $ABCD$.

For a given cell taken with an assigned direction, we define $\Delta x \Delta y \Delta z$ as the product obtained by multiplying the increments of x, y, z which are found by passing along consecutive edges of the cell in the assigned direction, first traversing an edge on which x alone varies, then one on which y alone varies, and finally one on which z alone varies. It is easy to verify that this fixes $\Delta x \Delta y \Delta z$ not only in magnitude, but also in sign when the cell is given and a direction assigned to it. Thus let the cell be a rectangular box $ABCDEFG$ formed by the parallel rectangles $ABCH$ and $FEDG$ (x varying on AB, y on BC, and z on CD). We may take Δx on AB, Δy on BC, and Δz on CB, or we may take Δx on EF, Δy on FG, and Δz on GH. In the second case two signs (those of Δx and Δz) are changed, and so $\Delta x \Delta y \Delta z$ is not changed.

Similarly, the product $\Delta y \Delta z \Delta x$ is obtained by traversing the edges of the cell in the same direction, first traversing an edge on which y alone varies, then one on which z alone varies, and finally, one on which x alone varies. The other products of Δx, Δy, Δz are defined in a similar way. It is easy to verify that each inversion of the order of Δx, Δy, Δz introduces a negative sign in the result. Thus

$$\Delta x \Delta y \Delta z = -\Delta x \Delta z \Delta y = \Delta z \Delta x \Delta y. \tag{3}$$

Let all the cells in a region Γ be taken in the same direction. If (x, y, z) is any point within the cell to which $\Delta x \Delta y \Delta z$ refers, the integral in the assigned direction over Γ,

$$\int\int\int f(x, y, z)\, dxdydz,$$

is defined as the limit (if it exists) approached by

$$\Sigma\Sigma\Sigma f(x, y, z) \Delta x \Delta y \Delta z$$

when Δx, Δy, and Δz approach zero, the summation being for all the regular cells within Γ. Similarly,

$$\int\int\int f(x, y, z)\, dydzdx = \text{Lim } \Sigma\Sigma\Sigma f(x, y, z) \Delta y \Delta z \Delta x,$$

etc. Hence, from (3),

$$\int\int\int f(x, y, z)\, dxdydz = -\int\int\int f(x, y, z)\, dxdzdy = \int\int\int f(x, y, z)\, dzdxdy. \tag{4}$$

As in case of double integration, one of the sets of surfaces $x=$const., $y=$const., $z=$const., can be replaced by a third set of surfaces, provided the irregular cells thus introduced have a total volume that approaches zero in the limit.

Expression of an Integral in Terms of Boundary Values.—A simple integral is expressed in terms of its limits by the formula

$$\int_a^b du = u \Big|_a^b = u(b) - u(a). \tag{5}$$

Similar formulas apply to multiple integrals.

Let u be a function of x and y. In the integral,

$$\int\int du dx,$$

let du be taken along $x=$const. We may take dx along $u=$const. or $y=$const. as we choose. The integral of du along a curve $x=$const. from one intersection with the border to another is given by equation (5). Hence, if we evaluate the integral by summing first with respect to u and then with respect to x, we get

$$\int\int du dx = \int u dx. \tag{6}$$

The double integral is taken over a region Γ, the simple integral over the boundary of Γ. Since du and dx occur on consecutive sides of a quadrilateral in the double integral, the direction of integration around the boundary must be such that if BC belongs to the boundary and $ABCD$ is a quadrilateral directed as in the double integral, then the integral along the boundary is in the direction BC.

Similarly, if u is a function of x, y, z,

$$\int\int\int du dx dy = \int\int u dx dy. \tag{7}$$

The triple integral is taken over a region Γ, the double integral over its boundary. The directions of integration are so related that if $ABCD$ is a cell of the triple integral with face BCD in the boundary, then BCD gives in that face the direction of the double integral.

Illustrations of these formulas are furnished by the theorems of Green, Stokes, and Gauss. Suppose, for example, P, Q, R are functions of x, y, z on a two-sided surface. Then

$$\int P dx = \int\int dP dx,$$

the two integrals being taken over a region of the surface and around its boundary, respectively. Since dP is determined on the curves $x=$const.,

$$dP = \frac{\partial P}{\partial y} dy + \frac{\partial P}{\partial z} dz,$$

and

$$\int P dx = \int\int \frac{\partial P}{\partial y} dy dx + \frac{\partial P}{\partial z} dz dx.$$

In these integrals dy and dz are taken along $x=\text{const.}$ In the combination $dydx$ we may take dx along $y=\text{const.}$ and in $dzdx$ along $z=\text{const.}$ Similarly,

$$\int Qdy=\int\int \frac{\partial Q}{\partial x}\,dxdy+\frac{\partial Q}{\partial z}\,dzdy,$$

$$\int Rdz=\int\int \frac{\partial R}{\partial x}\,dxdz+\frac{\partial R}{\partial y}\,dydz.$$

Adding these equations and changing the sign each time we invert the order of differentials, we get

$$\int Pdx+Qdy+Rdz=\int\int\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dxdy+\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)dydz+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)dzdx,$$

which is Stokes' Theorem. Since

$$dP=\frac{\partial P}{\partial x}\,dx+\frac{\partial P}{\partial y}\,dy+\frac{\partial P}{\partial z}\,dz,$$

the equation

$$\int\int dPdx=\int\int \frac{\partial P}{\partial y}\,dydx+\frac{\partial P}{\partial z}\,dzdx,$$

suggests that symbolically

$$\int\int \frac{\partial P}{\partial x}\,dxdx=0. \tag{8}$$

This is in line with our definition since one set of curves does not give rise to any regular quadrilaterals, and so the summation from which the integral might be defined is zero.

Change of Variable.—Let x, y be functions of u, v. Then

$$dx=\frac{\partial x}{\partial u}\,du+\frac{\partial x}{\partial v}\,dv$$

along any curve. Hence

$$\int\int fdxdy=\int\int f\frac{\partial x}{\partial u}\,dudy+\int\int f\frac{\partial x}{\partial v}\,dvdy.$$

In these integrals dy is determined along the curves $x=\text{const.}$ Since this is not significant we may determine dy in the first integral on the curves $v=\text{const.}$, and in the second integral on the curves $u=\text{const.}$ In the first case

$$dy=\frac{\partial y}{\partial v}\,dv,$$

and in the second

$$dy=\frac{\partial y}{\partial u}\,du,$$

whence

$$\int\int fdxdy=\int\int f\frac{\partial x}{\partial u}\frac{\partial y}{\partial v}\,dudv+\int\int f\frac{\partial x}{\partial v}\frac{\partial y}{\partial u}\,dvdu=\int\int f\left[\frac{\partial x}{\partial u}\frac{\partial y}{\partial v}-\frac{\partial x}{\partial v}\frac{\partial y}{\partial u}\right]dudv. \tag{9}$$

This result could be obtained by using the values

$$dx = \frac{\partial x}{\partial u} du + \frac{\partial x}{\partial v} dv,$$

$$dy = \frac{\partial y}{\partial u} du + \frac{\partial y}{\partial v} dv,$$

in $\int\int f dxdy$, expanding, and assuming that the integrals containing $dudu$ and $dvdv$ are zero.

In a similar way we show

$$\int\int\int f dxdydz = \int\int\int f \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \\ \frac{\partial x}{\partial w} & \frac{\partial y}{\partial w} & \frac{\partial z}{\partial w} \end{vmatrix} dudvdw. \qquad (10)$$

Equations (9) and (10) are valid whether the Jacobian has an invariable sign or not. Irregular cells will, however, usually occur in the neighborhood of a point where the Jacobian is zero. It may therefore be necessary that it be possible to enclose such points in an area in case of a double integral and in a volume in case of a triple integral which has a limit zero.

While an integral is represented by a number, the summation process upon which the integral is based is as much geometric as arithmetic. Symbolically the integral is a function in which $dxdy$ is equivalent to $-dydx$ and dx^2 to zero. This can be expressed by means of vectors. Yet dx, dy, dz are not vectors, and integration belongs no more to vector analysis than algebra does. They both belong to the larger field of quantities having sign, but not direction.

P-way Determinants, with an Application to Transvectants.

BY LEPINE HALL RICE.

In this paper an extended definition of a determinant is given which applies to determinants of more than three dimensions, and enables us to remove the restriction in Cayley's law of multiplication and to set up a new case in Scott's law of multiplication. New formulas are obtained for the known process of decomposition of a determinant into determinants of fewer dimensions, and a new process called crossed decomposition is described. Fresh light is thrown upon the function known as a "determinant-permanent," a limitation hitherto thought necessary being done away. Finally a generalization to p dimensions is made of Metzler's theorem in two dimensions concerning a determinant, each of whose elements is the product of k factors.

We lead up to these matters by a brief statement of the elementary theory of 3-way or cubic determinants and permanents.

I. THREE-WAY DETERMINANTS AND PERMANENTS.

1. *Definitions and Fundamental Properties.*—Elements, n^3 in number, can be arranged in a 3-way matrix of order n having n^2 *rows*, n^2 *columns*, and n^2 *normals;* all are called *files*. The matrix divides up into n *strata*, or layers that contain rows and columns and are pierced by normals; and into n *row-normal layers* pierced by columns; and into n *column-normal layers* pierced by rows; all are called *layers*. In triple-index notation, a_{ijk} denotes the element in the i-th stratum, the j-th row-normal layer, and the k-th column-normal layer. To represent 3-way matrices of successive orders, we write:

$$\left\| \begin{array}{cc|cc} a_{111} & a_{112} & a_{211} & a_{212} \\ a_{121} & a_{122} & a_{221} & a_{222} \end{array} \right\|, \quad \left\| \begin{array}{ccc|ccc|ccc} a_{111} & a_{112} & a_{113} & a_{211} & a_{212} & a_{213} & a_{311} & a_{312} & a_{313} \\ a_{121} & a_{122} & a_{123} & a_{221} & a_{222} & a_{223} & a_{321} & a_{322} & a_{323} \\ a_{131} & a_{132} & a_{133} & a_{231} & a_{232} & a_{233} & a_{331} & a_{332} & a_{333} \end{array} \right\|, \text{ etc.}$$

Two or more elements are *conjunctive* if no two of them lie in the same layer of any direction; n conjunctive elements are *perjunctive* and form a *transversal.* When we speak of a determinant of a matrix and use the word "transversal," we mean the product of these elements. The *locant* of an

element is the set of indices which locate it in the matrix. If the locants of a perjunctive set of elements be written in a column, the three subcolumns, of n indices each, are called *ranges*, and the whole is called the locant of the set (transversal); *e. g.*, $a_{122}\,a_{313}\,a_{231}$, ranges 132, 213, 231. The *sign of a range* is $+$ or $-$ according as there is an even or an odd number of inversions of order in the range.

The *determinant* of a 3-way matrix is the algebraic sum of its transversals, each having the sign which is the product of the signs of its second and third ranges.

The *permanent* of a 3-way matrix is the sum of its transversals.

The determinant and permanent being homogeneous linear functions of the elements of any layer, we have obvious theorems as to factors of layers, and as to separation into a sum of 3-way determinants or permanents when there are polynomial elements, and, conversely, as to addition of determinants and of permanents. An interchange of strata in a determinant, or of any parallel layers in a permanent, does not change its value; but an interchange of any two parallel layers other than strata, in a determinant, changes its sign. Hence, if two such layers are alike the determinant vanishes. Hence a multiple of such a layer may be added to any parallel layer without changing the value of the determinant.

A *minor* is formed by striking out an equal number of layers of each direction. It is obvious what we mean by *conjunctive minors* and by *perjunctive minors.* A perjunctive set of minors, formed into a product, with the proper sign, is equal to the sum of a certain number of terms of the determinant; the sign is the sign of that term of the determinant whose elements are the elements in the main diagonals of the minors. In the simple case of an element a_{ijk} and its complementary minor, the sign is $(-1)^{j+k}$. In consequence, the Laplacean expansion of a determinant is formed by partitioning the matrix into two or more sets of parallel layers and forming all possible perjunctive sets of minors occupying the sets of layers, one minor in each set.

2. *Decomposition.*—(i) A 3-way determinant Δ of order n can be decomposed into the sum of $n!$ 2-way determinants whose rows are rows of Δ. For, arranging n perjunctive rows of Δ in the order of the row-normal layers in which they lie, we see that the 2-way determinant

$$\begin{vmatrix} a_{i'11} & a_{i'12} & \dots\dots & a_{i'1n} \\ a_{i''21} & a_{i''22} & \dots\dots & a_{i''2n} \\ \dots & \dots & \dots\dots & \dots \\ a_{i^{(n)}n1} & a_{i^{(n)}n2} & \dots\dots & a_{i^{(n)}nn} \end{vmatrix},$$

whose matrix they form consists of $n!$ terms of Δ, and that the totality of such 2-way determinants consists of all the terms of Δ. They are called *components of* Δ.

(ii) If the rows are arranged in the order of the *strata* in which they lie, then each determinant must have prefixed the sign of the j-range (denoted by $\pm_j$) in the locant of the set of rows.

(iii) (iv) There will be two corresponding forms of decomposition into 2-way determinants whose columns are columns of Δ.

(v) We can also decompose Δ into an algebraic sum of 2-way permanents. Arrange n perjunctive normals in any order, and to the permanent whose matrix they form, prefix the sign of their locant; for, clearly, all of those $n!$ terms of Δ which lie in a perjunctive set of normals have the same sign.

$$\text{(ii)}\ \pm_j \begin{vmatrix} a_{1j'1} & a_{1j'2} & \cdots \\ a_{2j''1} & a_{2j''2} & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix}; \quad \text{(iii)} \begin{vmatrix} a_{i'11} & a_{i''12} & \cdots \\ a_{i'21} & a_{i''22} & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix};$$

$$\text{(iv)}\ \pm_k \begin{vmatrix} a_{11k'} & a_{21k''} & \cdots \\ a_{12k'} & a_{22k''} & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix}; \quad \text{(v)}\ \pm_j \pm_k \overset{+}{\Bigg|} \begin{matrix} a_{1j'k'} & a_{2j'k'} & \cdots \\ a_{1j''k''} & a_{2j''k''} & \cdots \\ \cdots & \cdots & \cdots \end{matrix} \overset{+}{\Bigg|}.^*$$

A 3-way *permanent* may of course be decomposed by rows, columns, or normals, into a sum of 2-way permanents.

3. *Element-Multiplication.*—(i) (Scott's† law of multiplication). The product of two 2-way determinants (or permanents) A and B of order n is expressible as a 3-way determinant (or permanent) C of order n wherein

$$c_{ijk} \equiv a_{ij} b_{ik}.$$

EXAMPLE:
$$\overset{+}{\Big|} \begin{matrix} a_{11}a_{12} \\ a_{21}a_{22} \end{matrix} \Big| \cdot \overset{+}{\Big|} \begin{matrix} b_{11}b_{12} \\ b_{21}b_{22} \end{matrix} \Big| = \overset{+}{\Big|} \begin{matrix} a_{11}b_{11} & a_{11}b_{12} \\ a_{12}b_{11} & a_{12}b_{12} \end{matrix} \Big| \begin{matrix} a_{21}b_{21} & a_{21}b_{22} \\ a_{22}b_{21} & a_{22}b_{22} \end{matrix} \Big|.$$

From the *prescription* $c_{ijk} \equiv a_{ij} b_{ik}$, whatever be the value of n, we see (i) noting the index i, that columns of A and columns of B are found in normals of C, (ii) noting the index j, that rows of A are found in columns of C and that elements of B run as factors through columns of C; and (iii) noting the index k, that rows of B are found in rows of C, and that elements of A run as factors through rows of C.

---

* "$\overset{+}{|} \ |$" means a determinant or permanent; "$\overset{+}{|} \ \overset{+}{|}$" means a permanent.

† R. F. Scott, "On Cubic Determinants," etc., *Proc. London Math. Soc.*, Vol. XI (1879), p. 17, at p. 23, paragraph 7. In paragraphs 8 and 9, Scott extends the rule so as to give the product of two determinants of p and q dimensions, respectively, in the form of a determinant of $p+q-1$ dimensions. In § 7 of the present paper this rule is extended to determinants as defined in § 5.

This method of examining a prescription will often give at once a good idea of a matrix and may suggest the best way of dealing with it.

In the present case we use decomposition (ii) of the preceding section and find that each component equals B multiplied by a term of A, giving $AB=C$.

(ii) The product of a 2-way determinant A and a 2-way permanent P of order n is expressible as a 3-way determinant C wherein

$$c_{ijk} \equiv a_{jk} p_{ik};$$

that is, a normal of C is a column of P with an a-factor; a column of C is a column of A with a p-factor; and a row of C is a row of A and a row of P. Using decomposition (v), we find that each component equals P multiplied by a term of A, and $AP=C$.

4. *File-Multiplication.*—(Cayley's* law of multiplication.) The product of a 3-way determinant A and a 2-way determinant B of order n is expressible as a 3-way determinant C of order n, wherein

$$c_{ijk} \equiv \sum_{l=1}^{n} a_{ijl} b_{kl};$$

that is, the matrix of each stratum of C is precisely what would result from using the familiar process of multiplying together two 2-way determinants into a 2-way determinant, row into row; one of the determinants being always B, and the other being the determinant of the matrix of the stratum of A corresponding to that of C. In brief, we multiply B into the strata of A to form C.

Using decomposition (i), we find that each component $C_r=A_rB$, whence $C=AB$.

II. Determinants and Permanents of p Dimensions.

5. *Definitions and Fundamental Properties.*—A p-way matrix (or a matrix of *class* p) of order n is formed of n^p elements:

$$\|a_{\alpha_1 \alpha_2 \ldots \alpha_p}\|_n^{(p)}.$$

The matrix can be separated into n layers, in any one of the p directions; a layer is a $(p-1)$-way matrix of n^{p-1} elements. Common to two layers of

* A. Cayley, "On the Theory of Determinants," *Trans. Cambridge Phil. Soc.*, Vol. VIII (1843), p. 75; *Coll. Math. Papers*, Vol. I, p. 63. See §2 of that paper. Cayley extends the rule so as to give the product of two determinants of p and q dimensions, respectively, in the form of a determinant of $p+q-2$ dimensions, but states that the rule is inapplicable when both p and q are odd; a restriction ignored by Scott in paragraph 9 of the paper cited in the first note, and by many others (see "Abrégé de la théorie des déterminants à n dimensions," par Maurice Lecat; Gand, Hoste, 1911, pp. 55 *et seq.*). In §8 of the present paper this rule is extended to determinants as defined in §5, such determinants serving to remove the restriction.

different directions is a $(p-2)$-way *sublayer;* common to three, a $(p-3)$-way sublayer; and so on, until we have, common to $p-1$ layers of different directions, a *file* of n elements, piercing the n parallel layers of the remaining direction. Parallel to this file are $n^{p-1}-1$ other files, together with it containing all the elements of the matrix. Files of the last direction are *rows.*

We speak of *conjunctive* and *perjunctive* elements, files, etc., of *transversals,* and of *locants* and *ranges,* as we did in the special case of a 3-way matrix.

Hitherto a p-way determinant has been defined as the algebraic sum of the transversals of a p-way matrix, the sign of a transversal being determined by arranging its elements in such an order that the values of a fixed index shall read $1, 2, \ldots, n$, and then taking the product of the signs of all the other ranges. It has then been shown that for a matrix of even class the same determinant will result, whatever be the fixed index, but that for a matrix of odd class a different determinant, in general, will result from a different choice of the fixed index. It has also been shown that interchange of layers denoted by the fixed index in a determinant of odd class leaves the value of the determinant unchanged, while the interchange of two layers of any other direction in such a determinant, or the interchange of two layers of any direction in a determinant of even class, changes the sign of the determinant.

It follows that without making the supposition that the elements of a transversal are first arranged in any particular order, we may say that the sign is the product (i) of the signs of all the ranges, in a determinant of even class, and (ii) of the signs of all but one (a fixed one) of the ranges, in a determinant of odd class.

From this point of view, we now generalize the definition of a determinant. We shall call an index (or range, or direction, or file, or set of layers) *signant* or *nonsignant* according as we do or do not take the order therein into account in fixing the sign of a term. In a 2-way determinant both indices are signant; in a 2-way permanent, both indices are nonsignant. In a 3-way determinant, two indices are signant. Passing to matrices of more dimensions than three, we see that it is possible not only to have signant all the indices if the class is even, and all but one if the class is odd, but to have signant a less number in either case, provided only that there be an even number that are signant—two indices in a 4-way or 5-way matrix, two or four indices in a 6-way or 7-way matrix, and so on. We therefore lay down the following:

Definition of a Determinant.—A determinant of a p-way matrix is the sum of all the terms that can be formed by taking a set of perjunctive elements

as factors and prefixing the product of the signs of an even number of chosen ranges.

A permanent might be viewed as one extreme, where the even number is zero; but file-multiplication and dependent processes have no application to permanents, and so we prefer to mention them explicitly when a theorem is true with regard to them, and to understand that a determinant has at least two signant indices.

By a *full-sign* determinant we shall mean a determinant as heretofore defined, with all, or all but one, of the indices signant, according as it is of even or odd class.

If in any determinant two layers of a signant direction be interchanged, the sign of the determinant is changed. Hence, if two such layers are alike, the determinant vanishes. Hence a multiple of one such layer may be added to another without changing the value of the determinant.

Both a determinant and a permanent are of course homogeneous linear functions of the elements of any layer and have the properties resulting.

6. *Decomposition.*—A p-way n-layer determinant Δ can be decomposed into the algebraic sum of $n!$ $(p-1)$-way determinants or permanents, as the case may be. Each of these components has $n!$ components, and so on. Ultimately we arrive at the expression of Δ as the algebraic sum of $(n!)^{p-2}$ 2-way determinants or permanents.

Let the matrix $\|a_{\alpha\beta\ldots\kappa\lambda}\|_n^{(p)}$ of Δ be divided up into its $(p-2)$-way sublayers of directions 1 and 2. Denote the sublayer common to the r-th layer of direction 1, and the s-th layer of direction 2 by $a_{rs0\ldots0}$. Take n perjunctive sublayers $a_{\alpha'10\ldots0}, a_{\alpha''20\ldots0}, \ldots, a_{\alpha^{(n)}n0\ldots0}$ to form a $(p-1)$-way matrix,

$$\left\| \begin{matrix} a_{\alpha'10\ldots0} \\ a_{\alpha''20\ldots0} \\ \cdots\cdots \\ a_{\alpha^{(n)}n0\ldots0} \end{matrix} \right\|.$$

This is a component of the matrix of Δ; there are $n!$ such components, and all of the locant of an element in Δ except the first index, is the locant of that element in each of the $(n-1)!$ components in which it occurs.

Denoting by the superscripts $\frown$ and $\smile$ the signant and nonsignant indices, and inserting a colon to isolate the locant of a component, we have:

$$\left.\begin{aligned}
&(1)\quad |a_{\frown{\alpha}\frown{\beta}\ldots}|_n^{(p)} = \sum_\alpha \pm_\alpha |a_{\alpha:\smile{\beta}\ldots}|_n^{(p-1)}; \\
&(2)\quad |a_{\frown{\alpha}\smile{\beta}\ldots}|_n^{(p)} = \sum_\alpha \pm_\alpha |a_{\alpha:\frown{\beta}\ldots}|_n^{(p-1)}; \\
&(3)\quad |a_{\smile{\alpha}\frown{\beta}\ldots}|_n^{(p)} = \sum_\alpha \quad |a_{\alpha:\frown{\beta}\ldots}|_n^{(p-1)}; \\
&(4)\quad |a_{\smile{\alpha}\smile{\beta}\ldots}|_n^{(p)} = \sum_\alpha \quad |a_{\alpha:\smile{\beta}\ldots}|_n^{(p-1)}.
\end{aligned}\right\} \qquad (D_1)$$

Each index beyond β is signant or nonsignant on the right according as it is signant or nonsignant on the left. In (1) and (2), $\pm_\alpha$ is the sign of the α-range in any transversal when the β-range reads $12\ldots.n$.

Briefly, if α is signant (nonsignant) the components are signed (unsigned) and the signancy of β is reversed (continued).

We see that the components of a determinant may be determinants or may be permanents; but that the components of a permanent must be permanents.

In verifying the formulas we must recall the fact that there are an even number of signant indices. Consider a term of $|a_{\alpha:\overset{\smile}{\overset{\frown}{\beta}}\ldots.}|_n^{(p-1)}$ in formula (1). Let its elements be arranged so that the values of β are in the order $12\ldots.n$; then its sign is the product of the signs of the signant ranges beyond β. Prefixing $\pm_\alpha$, we find that we now have the sign proper to this term in $|a_{\overset{\frown}{\alpha}\overset{\frown}{\beta}\ldots.}|_n^{(p)}$. And this sign is preserved when the elements are permuted, provided that we make β nonsignant, because there are an even number of signant ranges beyond β.

If the process be repeated to give the $(p-2)$-way components, we shall have $(n!)^2$ matrices of the form

$$\left\Vert \begin{matrix} a_{\alpha'\beta'10\ldots.0} \\ a_{\alpha''\beta''20\ldots.0} \\ \cdots\cdots\cdot \\ a_{\alpha^{(n)}\beta^{(n)}n0\ldots.0} \end{matrix} \right\Vert .$$

To condense the corresponding formulas, we use a double superfix $\overset{\smile}{\frown}$ or $\overset{\frown}{\smile}$, the upper signs to be read together and the lower signs together in each formula:

$$\left.\begin{aligned}
&(1)\quad |a_{\overset{\frown}{\alpha}\overset{\frown}{\beta}\overset{\smile}{\overset{\frown}{\gamma}}\ldots.}|_n^{(p)} = \sum_{\alpha,\beta} \pm_\alpha \pm_\beta |a_{\alpha\beta:\overset{\smile}{\overset{\frown}{\gamma}}\ldots.}|_n^{(p-2)}; \\
&(2)\quad |a_{\overset{\frown}{\alpha}\overset{\smile}{\beta}\overset{\smile}{\overset{\frown}{\gamma}}\ldots.}|_n^{(p)} = \sum_{\alpha,\beta} \pm_\alpha \qquad |a_{\alpha\beta:\overset{\circ}{\gamma}\ldots.}|_n^{(p-2)}; \\
&(3)\quad |a_{\overset{\smile}{\alpha}\overset{\frown}{\beta}\overset{\smile}{\overset{\frown}{\gamma}}\ldots.}|_n^{(p)} = \sum_{\alpha,\beta} \qquad \pm_\beta |a_{\alpha\beta:\overset{\circ}{\gamma}\ldots.}|_n^{(p-2)}; \\
&(4)\quad |a_{\overset{\smile}{\alpha}\overset{\smile}{\beta}\overset{\smile}{\overset{\frown}{\gamma}}\ldots.}|_n^{(p)} = \sum_{\alpha,\beta} \qquad\quad |a_{\alpha\beta:\overset{\smile}{\overset{\frown}{\gamma}}\ldots.}|_n^{(p-2)}.
\end{aligned}\right\} \qquad (D_2)$$

The formulas for *complete decomposition*, that is, separation into 2-way components, are:

$$\left.\begin{aligned}
&(1)\quad |a_{\ldots.\overset{\frown}{\kappa}\overset{\frown}{\lambda}}|_n^{(p)} = \Sigma \pm |a_{\ldots.:\overset{\frown}{\kappa}\overset{\frown}{\lambda}}|_n^{(2)}; \\
&(2)\quad |a_{\ldots.\overset{\smile}{\kappa}\overset{\frown}{\lambda}}|_n^{(p)} = \Sigma \pm |a_{\ldots.:\overset{\frown}{\kappa}\overset{\frown}{\lambda}}|_n^{(2)}; \\
&(3)\quad |a_{\ldots.\overset{\frown}{\kappa}\overset{\smile}{\lambda}}|_n^{(p)} = \Sigma \pm \overset{+}{|} a_{\ldots.:\overset{\smile}{\kappa}\overset{\smile}{\lambda}} \overset{+}{|}_n^{(2)}; \\
&(4)\quad |a_{\ldots.\overset{\smile}{\kappa}\overset{\smile}{\lambda}}|_n^{(p)} = \Sigma \pm \overset{+}{|} a_{\ldots.:\overset{\smile}{\kappa}\overset{\smile}{\lambda}} \overset{+}{|}_n^{(2)}.
\end{aligned}\right\} \qquad (D_{p-2})$$

The "$\pm$" is the product of the signs of all the signant ranges before κ.

It will be seen that the 2-way determinants in (1) and (2) and the 2-way permanents in (3) and (4) have for their rows files of Δ of the p-th direction (index λ) *i. e.*, rows of Δ; and that we get determinants when λ is signant, and permanents when λ is nonsignant.

As we can put the indices of any determinant in any order before decomposing it, the formulas are general with respect to such order. This remark applies to some of the later formulas.

In general, if we wish to have r nonsignant indices $\alpha_1\alpha_2\ldots\alpha_r$, and s signant indices $\beta_1\beta_2\ldots\beta_s$ come before the colon, an index γ to come immediately after it (that is, to be the index of the range which is the *base* to which $\pm_{\beta_1}, \pm_{\beta_2}, \ldots, \pm_{\beta_s}$ relate), and wish t nonsignant indices $\delta_1\delta_2\ldots\delta_t$ and u signant indices $\varepsilon_1\varepsilon_2\ldots\varepsilon_u$ to follow γ, the result is:

$$\left.\begin{array}{l} \left| a_{\overset{\smile}{\alpha_1}\ldots\overset{\smile}{\alpha_r}\overset{\frown}{\beta_1}\ldots\overset{\frown}{\beta_s}\overset{\smile}{\overset{\frown}{\gamma}}\overset{\smile}{\delta_1}\ldots\overset{\smile}{\delta_t}\overset{\frown}{\varepsilon_1}\ldots\overset{\frown}{\varepsilon_u}} \right|_n^{(r+s+t+u+1)} \\ \quad = \sum\limits_{\alpha_1\ldots\alpha_r\beta_1\ldots\beta_s} \pm_{\beta_1}\pm_{\beta_2}\ldots\pm_{\beta_s} \left| a_{\alpha_1\ldots\alpha_r\beta_1\ldots\beta_s:\overset{\frown^u}{\gamma}\overset{\smile}{\delta_1}\ldots\overset{\smile}{\delta_t}\overset{\frown}{\varepsilon_1}\ldots\overset{\frown}{\varepsilon_u}} \right|_n^{(t+u+1)}, \end{array}\right\} \quad (D)$$

where $\frown^u$ is to mean $\frown$ or $\smile$ according as u is odd or even, regardless of whether γ was originally signant or nonsignant. That is, the signancy of γ on the right is to be so taken that there will be an even number of signant indices in the components.

There is an interesting resemblance between the behavior of the signs $\smile$ and $\frown$ and that of $+$ and $-$. See (D_1) and (D_2), where, symbolically, $\frown\frown=\smile$, $\frown\smile=\frown$, $\smile\frown=\frown$, $\smile\smile=\smile$, $\frown\frown\overset{\smile}{\frown}=\overset{\smile}{\frown}$, etc. And, generally, in (D), taking the superfixes of the α's, the β's and γ, we find that $\smile\ldots\smile\frown\ldots\frown\overset{\smile}{\frown}=\overset{\smile}{\frown}$ if s is even, and $=\overset{\frown}{\smile}$ if s is odd.

7. *Element-Multiplication.*—If from the elements of the matrices

$$\| a_{\alpha_1\alpha_2\ldots\alpha_p} \|_n^{(p)}, \quad \| b_{\beta_1\beta_2\ldots\beta_q} \|_n^{(q)}$$

we form a third matrix of class $p+q-1$ and order n, in whose rows the rows of the a-matrix and the rows of the b-matrix are found according to the prescription

$$c_{\alpha_1\ldots\alpha_{p-1}\beta_1\ldots\beta_{q-1}\mu} \equiv a_{\alpha_1\ldots\alpha_{p-1}\mu}\, b_{\beta_1\ldots\beta_{q-1}\mu},$$

then it is seen that any transversal of the c-matrix consists of a transversal of the a-matrix and a transversal of the b-matrix, every possible combination occurring just once.

Let A be either a determinant or the permanent of the a-matrix, and let B be a determinant or the permanent of the b-matrix. Let C be a determinant or the permanent of the c-matrix, according to the result when the signancy of

$\alpha_1 \ldots \alpha_{p-1}$ in A, and of $\beta_1 \ldots \beta_{q-1}$ in B is continued in C, and when μ is made nonsignant in C if it is signant or is nonsignant in both A and B, but otherwise is made signant in C:

$$\frown\frown=\smile;\quad \smile\smile=\smile;\quad \smile\frown=\frown;\quad \frown\smile=\frown.$$

Then the evident theorem is:

$$AB=C.$$

The theorem includes as special cases both of the theorems of Section 3.

In the case $\frown\frown=\smile$, if either A or B is of odd class, C has more than one nonsignant index. Thus determinants that are not full-sign determinants not only fit into the cases previously known, but also create a new case.

8. *File-Multiplication.*—Given any two determinants with signant rows:

$$A \equiv \left| a_{\breve{\alpha}_1 \ldots \breve{\alpha}_f \overset{\frown}{\alpha}_{f+1} \ldots \overset{\frown}{\alpha}_p} \right|_n^{(p)}, \quad B \equiv \left| b_{\breve{\beta}_1 \ldots \breve{\beta}_g \overset{\frown}{\beta}_{g+1} \ldots \overset{\frown}{\beta}_q} \right|_n^{(q)},$$

let us compound every row of A into every row of B in the way familiar in the case of 2-way determinants and used in Section 4, so as to form a determinant C of class $p+q-2$ and order n, according to the prescription

$$c_{\breve{\alpha}_1 \ldots \breve{\alpha}_f \overset{\frown}{\alpha}_{f+1} \ldots \overset{\frown}{\alpha}_{p-1} \breve{\beta}_1 \ldots \breve{\beta}_g \overset{\frown}{\beta}_{g+1} \ldots \overset{\frown}{\beta}_{q-1}} \equiv \sum_{\mu=1}^{n} a_{\alpha_1 \ldots \alpha_{p-1}\mu} b_{\beta_1 \ldots \beta_{q-1}\mu};$$

that is, combine the locants of the rows of A and B to form the locant of an element of C, and continue the signancy of the indices in those locants. Then

$$AB=C.$$

PROOF: Completely decompose A and B into:

$$\Sigma \pm_{\alpha_{f+1}} \ldots \pm_{\alpha_{p-2}} \left| a_{\alpha_1 \ldots \alpha_{p-2} : \overset{\frown}{\alpha}_{p-1} \overset{\frown}{\alpha}_p} \right|_n^{(2)},$$

$$\Sigma \pm_{\beta_{g+1}} \ldots \pm_{\beta_{q-2}} \left| b_{\beta_1 \ldots \beta_{q-2} : \overset{\frown}{\beta}_{q-1} \overset{\frown}{\beta}_q} \right|_n^{(2)};$$

$$\text{or,}\quad \Sigma \pm_\alpha \begin{vmatrix} a_{\alpha'_1 \ldots \alpha'_{p-2} 11} & a_{\alpha'_1 \ldots \alpha'_{p-2} 12} & \ldots \\ a_{\alpha''_1 \ldots \alpha''_{p-2} 21} & a_{\alpha''_1 \ldots \alpha''_{p-2} 22} & \ldots \\ \ldots & \ldots & \ldots \end{vmatrix}_n, \quad \Sigma \pm_\beta \begin{vmatrix} b_{\beta'_1 \ldots \beta'_{q-2} 11} & b_{\beta'_1 \ldots \beta'_{q-2} 12} & \ldots \\ b_{\beta''_1 \ldots \beta''_{q-2} 21} & b_{\beta''_1 \ldots \beta''_{q-2} 22} & \ldots \\ \ldots & \ldots & \ldots \end{vmatrix}_n.$$

Multiply rowwise:

$$AB = \Sigma \pm_\alpha \pm_\beta \begin{vmatrix} \sum_\mu a_{\alpha'_1 \ldots \alpha'_{p-2} 1\mu} b_{\beta'_1 \ldots \beta'_{q-2} 1\mu} & \sum_\mu a_{\alpha'_1 \ldots \alpha'_{p-2} 1\mu} b_{\beta''_1 \ldots \beta''_{q-2} 2\mu} & \ldots \\ \sum_\mu a_{\alpha''_1 \ldots \alpha''_{p-2} 2\mu} b_{\beta'_1 \ldots \beta'_{q-2} 1\mu} & \sum_\mu a_{\alpha''_1 \ldots \alpha''_{p-2} 2\mu} b_{\beta''_1 \ldots \beta''_{q-2} 2\mu} & \ldots \\ \ldots & \ldots & \ldots \end{vmatrix}_n$$

$$= \Sigma \pm_\alpha \pm_\beta \begin{vmatrix} c_{\alpha'_1 \ldots \alpha'_{p-2} 1 \beta'_1 \ldots \beta'_{q-2} 1} & c_{\alpha'_1 \ldots \alpha'_{p-2} 1 \beta''_1 \ldots \beta''_{q-2} 2} & \ldots \\ c_{\alpha''_1 \ldots \alpha''_{p-2} 2 \beta'_1 \ldots \beta'_{q-2} 1} & c_{\alpha''_1 \ldots \alpha''_{p-2} 2 \beta''_1 \ldots \beta''_{q-2} 2} & \ldots \\ \ldots & \ldots & \ldots \end{vmatrix}_n.$$

The determinants in this sum are not components of C; but the $n!$ transversals of any one of them are $n!$ of the transversals of C, and the $(n!)^{p+q-4} \cdot n!$

transversals of all are the $(n!)^{p+q-3}$ transversals of C. As to sign, a transversal

$$c_{\alpha'_1 \ldots \alpha'_{p-2} 1 \beta_1^{(r_1)} \ldots \beta_{q-2}^{(r_1)} r_1} \; c_{\alpha''_1 \ldots \alpha''_{p-2} 2 \beta_1^{(r_2)} \ldots \beta_{q-2}^{(r_2)} r_2} \ldots$$

has the sign $\pm_\alpha \pm_{\beta^{(r)}} \pm_r$ in C, and the sign $\pm_\alpha \pm_\beta \pm_r$ in this sum; and since there are an even number of signant indices in $\beta_{1} \ldots \beta_{q-2}$, the signs $\pm_{\beta^{(r)}}$ and $\pm_\beta$ are alike. Therefore $AB = C$.

EXAMPLE: $A \equiv |a_{\breve{\alpha}_1 \hat{\alpha}_2 \hat{\alpha}_3}|_2^{(3)}$, $B \equiv |b_{\breve{\beta}_1 \hat{\beta}_2 \hat{\beta}_3 \hat{\beta}_4 \hat{\beta}_5}|_2^{(5)}$;

$$c_{\breve{\alpha}_1 \hat{\alpha}_2 \breve{\beta}_1 \hat{\beta}_2 \hat{\beta}_3 \hat{\beta}_4} \equiv a_{\alpha_1 \alpha_2 1} b_{\beta_1 \beta_2 \beta_3 \beta_4 1} + a_{\alpha_1 \alpha_2 2} b_{\beta_1 \beta_2 \beta_3 \beta_4 2}.$$

(It will be convenient here and elsewhere to insert commas in locants to bring out the structure of a matrix; they may be inserted, shifted or removed at pleasure, as they do not change the meaning of a locant.)

$$A = \begin{vmatrix} a_{1,11} a_{1,12} \\ a_{2,21} a_{2,22} \end{vmatrix} + \begin{vmatrix} a_{2,11} a_{2,12} \\ a_{1,21} a_{1,22} \end{vmatrix}, \quad B = \begin{vmatrix} b_{111,11} b_{111,12} \\ b_{222,21} b_{222,22} \end{vmatrix} - \begin{vmatrix} b_{112,11} b_{112,12} \\ b_{221,21} b_{221,22} \end{vmatrix} - \ldots;$$

$$AB = \begin{vmatrix} a_{11,1} b_{1111,1} + a_{11,2} b_{1111,2} & a_{11,1} b_{2222,1} + a_{11,2} b_{2222,2} \\ a_{22,1} b_{1111,1} + a_{22,2} b_{1111,2} & a_{22,1} b_{2222,1} + a_{22,2} b_{2222,2} \end{vmatrix}$$

$$- \begin{vmatrix} a_{11,1} b_{1121,1} + a_{11,2} b_{1121,2} & a_{11,1} b_{2212,1} + a_{11,2} b_{2212,2} \\ a_{22,1} b_{1121,1} + a_{22,2} b_{1121,2} & a_{22,1} b_{2212,1} + a_{22,2} b_{2212,2} \end{vmatrix} - \ldots$$

$$= \begin{vmatrix} c_{11,1111} c_{11,2222} \\ c_{22,1111} c_{22,2222} \end{vmatrix} - \begin{vmatrix} c_{11,1121} c_{11,2212} \\ c_{22,1121} c_{22,2212} \end{vmatrix} - \ldots = C.$$

This proof depends on the rows of A and B being signant. Apart from the proof, it is plain that one of the two indices $\alpha_p \beta_q$ can not be signant and the other nonsignant, for that would give C an odd number of signant indices. And we proceed to show that both indices can not be nonsignant; from which it follows that a full-sign determinant will not serve to express the file-product of two determinants of odd class—the restriction stated by Cayley in announcing his law of multiplication.

For, take $n=2$, to simplify the statement, In the transversal

$$c_{\alpha'_1 \ldots \alpha'_{p-2} 1 \, \beta'_1 \ldots \beta'_{q-2} 1} \; c_{\alpha''_1 \ldots \alpha''_{p-2} 2 \, \beta''_1 \ldots \beta''_{q-2} 2},$$

we find the monomial

$$a_{\alpha'_1 \ldots \alpha'_{p-2} 12} \; b_{\beta'_1 \ldots \beta'_{q-2} 12} \; a_{\alpha''_1 \ldots \alpha''_{p-2} 22} \; b_{\beta''_1 \ldots \beta''_{q-2} 22},$$

which does not consist of transversals of A and B, and we also find this monomial in the transversal

$$c_{\alpha'_1 \ldots \alpha'_{p-2} 1 \, \beta''_1 \ldots \beta''_{q-2} 2} \; c_{\alpha''_1 \ldots \alpha''_{p-2} 2 \, \beta'_1 \ldots \beta'_{q-2} 1};$$

and these two transversals are affected with the same sign, if α_p and β_q be assumed to be nonsignant, since there are then an even number of signant indices in $\alpha_1 \ldots . \alpha_{p-1}$ and in $\beta_1 \ldots . \beta_{q-1}$. Under this assumption, therefore, $AB \neq C$. The reasoning applies to every order of determinant and to every such monomial; if α_p and β_q in that monomial take one value v_1 times, and another value v_2 times, and so on, with $v_1 + v_2 + \ldots . = n$, then we shall find that monomial in $v_1! \, v_2! \ldots .$ transversals having the same sign prefixed.

It is instructive to compare the matrices resulting from element- and file-multiplication. If the polynomial elements of the latter matrix be converted into nonsignant files by deleting the $+$ signs, we have the former matrix, as is shown by the prescriptions.

9. *Crossed Decomposition.*—Let us generalize, in form and in substance, the development of a determinant which is found in the proof of the law of file-multiplication in the previous section. Beginning with the decomposition of any determinant or permanent

$$A \equiv | a_{\alpha_1 \ldots . \alpha_p} |_n^{(p)}$$

into

$$\Sigma \pm | a_{\alpha_1 \ldots . \alpha_{p-2} : \alpha_{p-1} \alpha_p} |_n^{(2)},$$

in each component alike let us cause any set of indices before the colon, provided that among them there are an even number of signant indices or else no signant indices, to take their n sets of values not in the n rows but in the n columns (as in the c-determinants in the above proof). Let us call the new determinants or permanents, with the signs of the components from which they were formed, *crossed components*. Example, $| a_{\breve{\alpha}_1 \hat{\alpha}_2 \hat{\alpha}_3 \hat{\alpha}_4 \hat{\alpha}_5} |_2^{(5)}$; a component, $\begin{vmatrix} a_{111,11} & a_{111,12} \\ a_{222,21} & a_{222,22} \end{vmatrix}$; the derived crossed component, $\begin{vmatrix} a_{111,11} & a_{122,12} \\ a_{211,21} & a_{222,22} \end{vmatrix}$.

The sum of these crossed components is equal to A. For, let the transposed indices be $\alpha_{h+1} \ldots . \alpha_{p-2}$, and consider any transversal of A:

$$a_{\alpha_1' \ldots . \alpha_{p-2}' 1 \alpha_p'} \, a_{\alpha_1'' \ldots . \alpha_{p-2}'' 2 \alpha_p''} \cdots a_{\alpha_1^{(n)} \ldots . \alpha_{p-2}^{(n)} n \alpha_p^{(n)}}.$$

Let $r_1 r_2 \ldots . r_n$ be such a permutation of $12 \ldots . n$ that

$$\alpha_p^{(r_1)} = 1, \; \alpha_p^{(r_2)} = 2, \; \ldots ., \; \alpha_p^{(r_n)} = n.$$

This transversal will be found once and only once, in and only in that crossed component whose main diagonal is

$$a_{\alpha_1' \ldots . \alpha_h' \alpha_{h+1}^{(r_1)} \ldots . \alpha_{p-2}^{(r_1)} 11} \, a_{\alpha_1'' \ldots . \alpha_h'' \alpha_{h+1}^{(r_2)} \ldots . \alpha_{p-2}^{(r_2)} 22} \cdots a_{\alpha_1^{(n)} \ldots . \alpha_h^{(n)} \alpha_{h+1}^{(r_n)} \ldots . \alpha_{p-2}^{(r_n)} nn},$$

i. e., in the crossed component

$$\pm \left| \begin{matrix} a_{\alpha_1' \ldots . \alpha_h' \alpha_{h+1}^{(r_1)} \ldots . \alpha_{p-2}^{(r_1)} 11} & a_{\alpha_1' \ldots . \alpha_h' \alpha_{h+1}^{(r_2)} \ldots . \alpha_{p-2}^{(r_2)} 12} & \ldots . \\ a_{\alpha_1'' \ldots . \alpha_h'' \alpha_{h+1}^{(r_1)} \ldots . \alpha_{p-2}^{(r_1)} 21} & a_{\alpha_1'' \ldots . \alpha_h'' \alpha_{h+1}^{(r_2)} \ldots . \alpha_{p-2}^{(r_2)} 22} & \ldots . \\ \ldots\ldots\ldots & \ldots\ldots\ldots & \end{matrix} \right|_n ;$$

and the sign will be right, for if by the symbol

$$\pm \widehat{\alpha}^{(r_1 \dots r_n)}_{h+1 \dots p-2}$$

we denote the product of the signs of the signant ranges among $\alpha_{h+1} \dots \alpha_{p-2}$ when they take their values in the order indicated by the symbols (r_1), (r_2), $\dots$, (r_n), then we have the equation

$$\pm \widehat{\alpha}^{(' \dots n)}_{1 \dots h} \pm \widehat{\alpha}^{(' \dots n)}_{h+1 \dots p-2} \pm \alpha^{(' \dots n)}_{p} = \pm \widehat{\alpha}^{(' \dots n)}_{1 \dots h} \pm \widehat{\alpha}^{(r_1 \dots r_n)}_{h+1 \dots p-2} \pm \alpha^{(' \dots n)}_{p}.$$

Generalizing in substance, let us decompose any determinant or permanent A into components of class q:

$$A = \Sigma \pm \left| a_{\alpha_{11} \dots \alpha_{1h_1} \alpha_{21} \dots \alpha_{2h_2} \dots \alpha_{q1} \dots \alpha_{qh_q} : \alpha_1 \alpha_2 \dots \alpha_q} \right|^{(q)}_{n},$$

the indices before the colon being separated into q groups with an even number of signant indices (0 even) in each group other than the first. Alter each component by causing each group other than the first to take its n sets of values not in the n 1st-way layers, but in the n s-th-way layers for the s-th group. By this alteration the main diagonal is unchanged. Retain the signs of the original components for these *crossed components;* in other words, give each crossed component the sign of its main diagonal term as a term of A.

The sum of these crossed components is equal to A.

EXAMPLE: The determinant $\left| a_{\breve{\alpha}_{11} \alpha_{12} \widehat{\alpha}_{21} \widehat{\alpha}_{22} \widehat{\alpha}_{31} \widehat{\alpha}_{32} \widehat{\alpha}_1 \widehat{\alpha}_2 \widehat{\alpha}_3} \right|^{(9)}_{2}$ has 3-way components,

$$\Sigma \pm_{\alpha_{12}} \pm_{\alpha_{21}} \pm_{\alpha_{22}} \pm_{\alpha_{31}} \pm_{\alpha_{32}} \left| a_{\alpha_{11} \alpha_{12} \alpha_{21} \alpha_{22} \alpha_{31} \alpha_{32} : \breve{\alpha}_1 \widehat{\alpha}_2 \widehat{\alpha}_3} \right|^{(3)}_{2}.$$

A component: $+ \left| \begin{matrix} a_{11,11,11,,111} & a_{11,11,11,,112} \\ a_{11,11,11,,121} & a_{11,11,11,,122} \end{matrix} \middle| \begin{matrix} a_{22,22,22,,211} & a_{22,22,22,,212} \\ a_{22,22,22,,221} & a_{22,22,22,,222} \end{matrix} \right|$; the corresponding crossed component:

$$+ \left| \begin{matrix} a_{11,11,11,,111} & a_{11,11,22,,112} \\ a_{11,22,11,,121} & a_{11,22,22,,122} \end{matrix} \middle| \begin{matrix} a_{22,11,11,,211} & a_{22,11,22,,212} \\ a_{22,22,11,,221} & a_{22,22,22,,222} \end{matrix} \right|.$$

PROOF: Let $\prod\limits_{t=1}^{n} a_{\alpha^{(t)}_{11} \dots \alpha^{(t)}_{1h_1} \alpha^{(t)}_{21} \dots \alpha^{(t)}_{2h_2} \dots \alpha^{(t)}_{q1} \dots \alpha^{(t)}_{qh_q} t \alpha^{(t)}_{2} \dots \alpha^{(t)}_{q}}$ be any transversal of A. Take $q-1$ permutations of $12 \dots n$, viz., $r_{s1} r_{s2} \dots r_{sn}$, $s = 2, 3, \dots, q$, such that

$$\alpha^{(r_{s1})}_{s} = 1, \ \alpha^{(r_{s2})}_{s} = 2, \ \dots, \ \alpha^{(r_{sn})}_{s} = n.$$

The transversal will be found once and only once, in and only in that crossed component whose main diagonal is:

$$\prod_{t=1}^{n} a \left(\alpha^{(t)}_{11} \dots \alpha^{(t)}_{1h_1} \alpha^{(r_{2t})}_{21} \dots \alpha^{(r_{2t})}_{2h_2} \dots \alpha^{(r_{qt})}_{q1} \dots \alpha^{(r_{qt})}_{qh_q} tt \dots t \right).$$

And the sign will be right, since, for each value of s,

$$\pm \widehat{\alpha}^{(' \dots n)}_{s1 \dots sh_s} = \pm \widehat{\alpha}^{(r_{s1} \dots r_{sn})}_{s1 \dots sh_s}.$$

Thus, in the example above, the transversal

$$a_{112211121}\, a_{221122212}$$

must be found, if at all, in the position

$$\left| \begin{matrix} \\ a_{11,22,11,,121} \end{matrix} \right| \begin{matrix} a_{22,11,22,,212} \\ \\ \end{matrix} \right|$$

in a crossed component; and there is evidently one and only one crossed component in which it is found.

The development in crossed components may be written:

$$A=\sum(\prod_s \pm \widehat{\alpha}_{s1\ldots\ldots sh_s}) \left| a \left\{ \begin{array}{c|c} \alpha_{11}\cdot\ldots\cdot\alpha_{1h_1} & \alpha_1 \\ \alpha_{21}\cdot\ldots\cdot\alpha_{2h_2} & \alpha_2 \\ \cdots\cdots\cdots & \cdot\cdot \\ \alpha_{q1}\cdot\ldots\cdot\alpha_{qh_q} & \alpha_q \end{array} \right\} \right|_n^{(q)}.$$

10. *Raising and Lowering the Class: The "Determinant-Permanent."*— One may arbitrarily raise the class of a given determinant or permanent by introducing one or more new indices whose values are determined by the values of one or more of the original indices, suitably adjusting their signancy. For example, using Kronecker's symbol $\delta_{i_1\ldots\ldots i_s}=1$ if $i_1=\ldots\ldots=i_s$, otherwise $=0$, we have:

$$|a_{\widehat{\beta}\widehat{\gamma}}|_2^{(2)}=|\delta_{\alpha\beta}a_{\breve{\alpha}\widehat{\beta}\widehat{\gamma}}|_2^{(3)};\ i.\ e.,\ \begin{vmatrix} a_{11}a_{12} \\ a_{21}a_{22} \end{vmatrix} = \left| \begin{matrix} a_{11,1}\, a_{11,2} \\ \cdot \quad \cdot \end{matrix} \right| \begin{matrix} \cdot \quad \cdot \\ a_{22,1}\, a_{22,2} \end{matrix} \right|.$$

And in general we may introduce as many nonsignant indices as we like, each having any one-to-one correspondence that we like with one of the original indices; and may then arbitrarily render signant any two of the entire set of indices which happen to have a one-to-one correspondence with each other such that both have the same sign.

In particular, if a determinant of even class have two nonsignant indices, there may be introduced a signant index which shall everywhere take the same values as one of the nonsignant indices, the latter being also made signant, and thus being *doubled*, the result being a determinant of odd class with one nonsignant index. *I. e.*,

$$|a_{\breve{\beta}\breve{\gamma}\ldots\ldots}|_n^{(2q)}=|\delta_{\alpha\beta}a_{\widehat{\alpha}\widehat{\beta}\breve{\gamma}\ldots\ldots}|_n^{(2q+1)}.$$

It is by this kind of a determinant of odd class that the product of two full-sign determinants $A^{(p)}$ and $B^{(q)}$ by file-multiplication has heretofore been expressed, when p and q were odd, as a determinant C of class $p+q-1$ (not $p+q-2$); the "fixed index" (nonsignant index) of A has been doubled in C,

and the "fixed index" of B has been made the "fixed index" of C. Of course C would consist largely of zeros.*

On the other hand, let us start with a full-sign determinant A, of odd class, of a type of which C in the last paragraph is a particular case; wherein a group of indices take the same values, another group take the same values, and so on, there being r groups with an even number of indices in each, s groups with an odd number in each, and t single indices, among the latter being the nonsignant index τ_t.† Give to the elements new locants by striking out all but one of each group of indices, put them into a matrix

$$\| a'_{\rho_1 \ldots \rho_r \sigma_1 \ldots \sigma_s \tau_1 \ldots \tau_t} \|_n^{(r+s+t)},$$

and consider the determinant

$$A' \equiv | a'_{\breve{\rho}_1 \ldots \breve{\rho}_r \overset{\frown}{\sigma}_1 \ldots \overset{\frown}{\sigma}_s \overset{\frown}{\tau}_1 \ldots \overset{\frown}{\tau}_{t-1} \breve{\tau}_t} |_n^{(r+s+t)};$$

noting that $s+t$ is necessarily odd. Evidently,

$$A' = A.$$

Now A' is not a full-sign determinant, but from A' there could be formed a function not involving anything but full-sign determinants and permanents, viz., the function invented by Gegenbauer† and called by him a "determinant-permanent." Decompose A', using (D) of Section 6, into

$$\Sigma \pm_{\sigma_1} \ldots \pm_{\sigma_s} \pm_{\tau_1} \ldots \pm_{\tau_{t-1}} | a'_{\sigma_1 \ldots \sigma_s \tau_1 \ldots \tau_{t-1} : \breve{\tau}_t \breve{\rho}_1 \ldots \breve{\rho}_r} |_n^{(r+1)}.$$

There are $(n!)^{s+t-1}$ of these permanents, and in them the $(\sigma_1 \ldots \sigma_s \tau_1 \ldots \tau_{t-1})$-ranges are differently written before the τ_t-range, and prefixed to each of them is the sign which is the product of the signs of these ranges. Such being the case, we can arbitrarily construct a determinant B with purely formal elements:

$$B \equiv | b_{\overset{\frown}{\sigma}_1 \ldots \overset{\frown}{\sigma}_s \overset{\frown}{\tau}_1 \ldots \overset{\frown}{\tau}_{t-1} \breve{\tau}_t} |_n^{(s+t)},$$

which will have the property that if for each transversal in its expansion in terms we substitute the corresponding permanent, the resulting function will be equal to A. This function is a "determinant-permanent" of class $s+t$ and *genus* $r+1$.

If A be made of *even* class by deleting τ_t, A' will be of class $r+s+t-1$, its components will be of class r, the indices after the colon will be $\breve{\rho}_1 \ldots \breve{\rho}_r$, and B will become

$$| b_{\overset{\frown}{\sigma}_1 \ldots \overset{\frown}{\sigma}_s \overset{\frown}{\tau}_1 \ldots \overset{\frown}{\tau}_{t-1} \breve{\rho}_1} |_n^{(s+t)};$$

---

* M. Lecat, "Sur la multiplication des déterminants," *Ann. Soc. Sci. de Bruxelles*, Vol. XXXVII, Part 2, p. 285.

† Abrégé, p. 32.

† L. Gegenbauer, "Einige Sätze über Determinanten höheren Ranges," *Denkschr. Akad. Wien*, Vol. LVII (1890), p. 735. Abrégé (see Note 3), p. 32.

that is, the "determinant-permanent" will be of class $s+t$ and genus r.

In either case the "determinant-permanent" is simply a decomposition of A'.

It has been said that a "determinant-permanent" is necessarily of odd class.* But the decomposition formula (D) shows us that there are two possible decompositions of a determinant $\left| a_{\overset{\frown}{\alpha_1}\cdots\overset{\frown}{\alpha_f}\breve{\alpha}_{f+1}\cdots\breve{\alpha}_p} \right|_n^{(p)}$ into a sum of permanents, such as to leave only signant indices before the colon, namely:

$$\Sigma \pm_{\alpha_1} \cdots \pm_{\alpha_f} \left| a_{\alpha_1 \ldots \alpha_f \,:\, \breve{\alpha}_{f+1} \ldots \breve{\alpha}_p} \right|_n^{(p-f)},$$
$$\Sigma \pm_{\alpha_1} \cdots \pm_{\alpha_{f-1}} \left| a_{\alpha_1 \ldots \alpha_{f-1} \,:\, \breve{\alpha}_f \ldots \breve{\alpha}_p} \right|_n^{(p-f+1)}.$$

Let A', therefore, be decomposed into

$$\Sigma \pm_{\sigma_1} \cdots \pm_{\sigma_s} \pm_{\tau_1} \cdots \pm_{\tau_{t-2}} \left| a'_{\sigma_1 \ldots \sigma_s \tau_1 \ldots \tau_{t-2} \,:\, \breve{\tau}_{t-1} (\breve{\tau}_t) \breve{\rho}_1 \ldots \breve{\rho}_r} \right|_n^{(r+1(+1))},$$

then

$$B \equiv \left| b_{\overset{\frown}{\sigma_1} \cdots \overset{\frown}{\sigma_s} \overset{\frown}{\tau_1} \cdots \overset{\frown}{\tau_{t-1}}} \right|_n^{(s+t-1)},$$

and the "determinant-permanent" is of *even* class.

EXAMPLE:

$$A \equiv \left| a_{\overset{\frown}{\alpha_1}\overset{\frown}{\beta_2}\overset{\frown}{\beta_3}\overset{\frown}{\beta_4}\overset{\frown}{\gamma_5}\overset{\frown}{\gamma_6}\overset{\frown}{\epsilon_7}\overset{\frown}{\epsilon_8}\overset{\frown}{\epsilon_9}\overset{\frown}{\epsilon_{10}}} \delta_{\beta_2\beta_3\beta_4} \delta_{\gamma_5\gamma_6} \delta_{\epsilon_7\epsilon_8\epsilon_9\epsilon_{10}} \right|_2^{(10)}.$$
$$A' \equiv \left| a'_{\overset{\frown}{\alpha}\overset{\frown}{\beta}\breve{\gamma}\breve{\epsilon}} \right|_2^{(4)} = (\mathrm{i})\, \Sigma \pm_\alpha \pm_\beta \left| a'_{\alpha\beta \,:\, \breve{\gamma}\breve{\epsilon}} \right|_2^{(2)} = (\mathrm{ii})\, \Sigma \pm_\alpha \left| a'_{\alpha \,:\, \breve{\beta}\breve{\gamma}\breve{\epsilon}} \right|_2^{(3)}.$$
$$B = (\mathrm{i}) \left| b_{\overset{\frown}{\alpha}\overset{\frown}{\beta}\breve{\gamma}} \right|_2^{(3)} \cdot B = (\mathrm{ii}) \left| b_{\overset{\frown}{\alpha}\overset{\frown}{\beta}} \right|_2^{(2)}.$$

(i) Replace $b_{111}b_{222} \qquad +b_{112}b_{221} \qquad -b_{121}b_{212} \qquad -b_{122}b_{211}$

by

$$\begin{vmatrix} a'_{1111} a'_{1112} \\ a'_{2221} a'_{2222} \end{vmatrix} + \begin{vmatrix} a'_{1121} a'_{1122} \\ a'_{2211} a'_{2212} \end{vmatrix} - \begin{vmatrix} a'_{1211} a'_{1212} \\ a'_{2121} a'_{2122} \end{vmatrix} - \begin{vmatrix} a'_{1221} a'_{1222} \\ a'_{2111} a'_{2112} \end{vmatrix},$$

that is, by

$$\begin{vmatrix} a_{1,111,11,1111} & a_{1,111,11,2222} \\ a_{2,222,22,1111} & a_{2,222,22,2222} \end{vmatrix} + \cdots.$$

(ii) Replace $\qquad b_{11}b_{22} \qquad -b_{12}b_{21}$

by

$$\begin{vmatrix} a'_{1111} a'_{1112} & a'_{2211} a'_{2212} \\ a'_{1121} a'_{1122} & a'_{2221} a'_{2222} \end{vmatrix} - \begin{vmatrix} a'_{2111} a'_{2112} & a'_{1211} a'_{1212} \\ a'_{2121} a'_{2122} & a'_{1221} a'_{1222} \end{vmatrix},$$

and translate into a's as in (i).

In the particular case of the determinant C described in this section, we have, first:

$$\left| a_{\breve{\alpha}_1\overset{\frown}{\alpha_2}\cdots\overset{\frown}{\alpha_p}} \right|_n^{(p)} \cdot \left| b_{\breve{\beta}_1\overset{\frown}{\beta_2}\cdots\overset{\frown}{\beta_q}} \right|_n^{(q)} = \left| c_{\overset{\frown}{\alpha_1}\overset{\frown}{\alpha_1}\overset{\frown}{\alpha_2}\cdots\overset{\frown}{\alpha_{p-1}}\breve{\beta}_1\overset{\frown}{\beta_2}\cdots\overset{\frown}{\beta_{q-1}}} \right|_n^{(p+q-1)}$$
$$= \Sigma \pm_{\alpha_2} \cdots \pm_{\alpha_{p-1}} \pm_{\beta_2} \cdots \pm_{\beta_{q-1}} \left| c'_{\alpha_2 \ldots \alpha_{p-1} \beta_2 \ldots \beta_{q-1} \,:\, \breve{\beta}_1 \breve{\alpha}_1} \right|_n^{(2)}, \quad (1)$$

which gives the product of A and B in the form of a "determinant-permanent" of class $p+q-3$ and genus 2, a form previously known. We attach the prime

* Abrégé, p. 32.

to c to give notice that one index of the group $\alpha_1\alpha_1$ has been dropped, just as a' was used above when all but one index in each group had been deleted.

And secondly we have the new form, obtained by writing after the colon any one of the indices that are before it in (1):

$$\Sigma \pm_{\alpha_2} \cdots \pm_{\alpha_{p-1}} \pm_{\beta_2} \cdots \pm_{\beta_{q-2}} \left| c'_{\alpha_2 \ldots \alpha_{p-1}\beta_2 \ldots \beta_{q-2} : \breve{\beta}_{q-1}\breve{\beta}_1\breve{\alpha}_1} \right|_n^{(3)}, \tag{2}$$

a "determinant-permanent" of even class $p+q-4$ and genus 3.

11. *A Product Determinant.*—A certain example given by Muir was generalized by Metzler in a paper "On a Determinant Each of Whose Elements is the Product of k Factors." * E. H. Moore contributed to the subject † and mentioned that a particular case of the theorem was ascribed to Kronecker. All this work was in two dimensions. Then von Sterneck extended Kronecker's result to p dimensions. ‡ We shall extend Metzler's theorem to p dimensions, including von Sterneck's generalization as a special case.

First let us deal with two sets of determinants, $A^{(1)}, A^{(2)}, \ldots, A^{(l)}$; $B^{(1)}, B^{(2)}, \ldots, B^{(k)}$:

$$A^{(\beta)} \equiv \left| a^{(\beta)}_{\breve{\alpha}_1 \ldots \breve{\alpha}_f \hat{\alpha}_{f+1} \ldots \hat{\alpha}_p} \right|_k^{(p)}; \quad B^{(\alpha)} \equiv \left| b^{(\alpha)}_{\breve{\beta}_1 \ldots \breve{\beta}_g \hat{\beta}_{g+1} \ldots \hat{\beta}_q} \right|_l^{(q)}.$$

Form a p-way determinant $\bar{A} = A^{(1)}A^{(2)} \ldots A^{(l)}$, of order kl, by placing $A^{(1)}, A^{(2)}, \ldots, A^{(l)}$ along the main diagonal, all other elements being zeros. Using the bipartite signs of order

$$11, 21, \ldots, k1, \quad 12, 22, \ldots, k2, \ldots, \quad 1l, 2l, \ldots, kl,$$

such as Moore employs, we shall have the prescription

$$\bar{a}_{(\breve{\alpha_1\beta})(\breve{\alpha_2\beta}) \ldots (\breve{\alpha_f\beta})(\widehat{\alpha_{f+1}\beta}) \ldots (\widehat{\alpha_p\beta})} \equiv a^{(\beta)}_{\alpha_1 \ldots \alpha_p}.$$

In the same way, form a q-way determinant of order kl, with $B^{(1)}, B^{(2)}, \ldots, B^{(k)}$; then alter its form by a rearrangement of layers, placing the l layers of the first direction containing $B^{(1)}$ in the positions $11, 12, \ldots, 1l$, those containing $B^{(2)}$ in the positions $21, 22, \ldots, 2l$, and so on, and do this for each direction. The resulting determinant $\bar{B} = B^{(1)}B^{(2)} \ldots B^{(k)}$ and the prescription is:

$$\bar{b}_{(\breve{\alpha\beta_1}) \ldots (\breve{\alpha\beta_g})(\widehat{\alpha\beta_{g+1}}) \ldots (\widehat{\alpha\beta_q})} \equiv b^{(\alpha)}_{\beta_1 \ldots \beta_q}.$$

---

* *Am. Math. Monthly*, Vol. VII (1900), p. 151.

† "A Fundamental Remark Concerning Determinantal Notations with the Evaluation of an Important Determinant of Special Form," *Annals of Math.*, Vol. (2) I, p. 177.

‡ R. D. von Sterneck, "Ausdehnung eines Kronecker'schen Satzes auf Determinanten höheren Ranges," *Rend. Palermo*, Vol. XXX (1910), p. 58. Lecat points out the fact that von Sterneck's theorem does not hold when the classes are both odd; Abrégé. p. 63. The present extension is not subject to that restriction.

Multiply together $\bar{A}$ and $\bar{B}$ by rows into a determinant U of class $p+q-2$ and order kl. The elements of U which are not zeros are monomials, since a row of $\bar{A}$ that is not blank contains nonzero elements only in the β-th set of k places, while a row of $\bar{B}$ that is not blank contains only one nonzero element in each set of k places. The a-element and b-element whose product is a u-element will be the a-element in whose locant $\alpha_p=\alpha$ and the b-element in whose locant $\beta_q=\beta$; we therefore change α to α_p and β to β_q to form the resulting prescription:

$$u_{\overset{\smile}{(\alpha_1\beta_q)}\overset{\cdots\cdot}{\cdots\cdot}\overset{\smile}{(\alpha_f\beta_q)}\overset{\frown}{(\alpha_{f+1}\beta_q)}\overset{\cdots\cdot}{\cdots\cdot}\overset{\frown}{(\alpha_{p-1}\beta_q)}\overset{\smile}{(\alpha_p\beta_1)}\overset{\cdots\cdot}{\cdots\cdot}\overset{\smile}{(\alpha_p\beta_g)}\overset{\frown}{(\alpha_p\beta_{g+1})}\overset{\cdots\cdot}{\cdots\cdot}\overset{\frown}{(\alpha_p\beta_{q-1})}}\equiv a^{(\beta_q)}_{\alpha_1\cdots\cdot\alpha_p}\,b^{(\alpha_p)}_{\beta_1\cdots\cdot\beta_q},$$

where at least $(\alpha_{p-1}\beta_q)$ and $(\alpha_p\beta_{q-1})$ are signant.

This gives the theorem

$$U=A^{(1)}\cdots\cdot A^{(l)}B^{(1)}\cdots\cdot B^{(k)}.$$

For 2-way determinants the prescription becomes

$$u_{\overset{\frown}{(\alpha_1\beta_2)}\overset{\frown}{(\alpha_2\beta_1)}}\equiv a^{(\beta_2)}_{\alpha_1\alpha_2}\,b^{(\alpha_2)}_{\beta_1\beta_2},$$

which agrees with the theorem designated T_2 by Moore.

If $A^{(1)}=A^{(2)}=\cdots\cdot=A^{(l)}=A$, say, and $B^{(1)}=B^{(2)}=\cdots\cdot=B^{(k)}=B$, we have:

$$U=A^lB^k,$$

which is an extension of von Sterneck's theorem to less than full-sign determinants.

EXAMPLE: $p=3$, $q=3$, $k=2$, $l=2$. $A'\equiv|a'_{\breve{\alpha}_1\widehat{\alpha}_2\widehat{\alpha}_3}|^{(3)}_2$, $A''\equiv|a''_{\breve{\alpha}_1\widehat{\alpha}_2\widehat{\alpha}_3}|^{(3)}_2$, $B'\equiv|b'_{\breve{\beta}_1\widehat{\beta}_2\widehat{\beta}_3}|^{(3)}_2$, $B''\equiv|b''_{\breve{\beta}_1\widehat{\beta}_2\widehat{\beta}_3}|^{(3)}_2$. For $\bar{A}$ and $\bar{B}$ we have:

$$\begin{array}{c|cccc|cccc|cccc|cccc|}
 & & (11) & & & & (21) & & & & (12) & & & & (22) & & \\
 & (11) & (21) & (12) & (22) & (11) & (21) & (12) & (22) & (11) & (21) & (12) & (22) & (11) & (21) & (12) & (22) \\
(11) & a'_{111} & a'_{112} & \cdot & \cdot & a'_{211} & a'_{212} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
(21) & a'_{121} & a'_{122} & \cdot & \cdot & a'_{221} & a'_{222} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
(12) & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & a''_{111} & a''_{112} & \cdot & \cdot & a''_{211} & a''_{212} \\
(22) & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & a''_{121} & a''_{122} & \cdot & \cdot & a''_{221} & a''_{222}
\end{array};$$

$$\begin{array}{c|cccc|cccc|cccc|cccc|}
(11) & b'_{111} & \cdot & b'_{112} & \cdot & \cdot & \cdot & \cdot & \cdot & b'_{211} & \cdot & b'_{212} & \cdot & \cdot & \cdot & \cdot & \cdot \\
(21) & \cdot & \cdot & \cdot & \cdot & \cdot & b''_{111} & \cdot & b''_{112} & \cdot & \cdot & \cdot & \cdot & \cdot & b''_{211} & \cdot & b''_{212} \\
(12) & b'_{121} & \cdot & b'_{122} & \cdot & \cdot & \cdot & \cdot & \cdot & b'_{221} & \cdot & b'_{222} & \cdot & \cdot & \cdot & \cdot & \cdot \\
(22) & \cdot & \cdot & \cdot & \cdot & \cdot & b''_{121} & \cdot & b''_{122} & \cdot & \cdot & \cdot & \cdot & \cdot & b''_{221} & \cdot & b''_{222}
\end{array}.$$

The prescription is:

$$u_{\overset{\smile}{(\alpha_1\beta_2)}\overset{\frown}{(\alpha_2\beta_2)}\overset{\smile}{(\alpha_3\beta_1)}\overset{\frown}{(\alpha_3\beta_2)}}\equiv a^{(\beta_3)}_{\alpha_1\alpha_2\alpha_3}\,b^{(\alpha_3)}_{\beta_1\beta_2\beta_3}.$$

We shall condense the representation of U by writing the b-factor of each element under the a-factor. It is convenient to have the four values of the first index of u appear in the four large horizontal subdivisions; those of the second in the large vertical subdivisions; and those of the third and fourth in the horizontal and vertical lines in each square common to two intersecting subdivisions. Thus one of the sixteen rows of $\overline{A}$ is associated with each of the sixteen squares, while one of the sixteen rows of $\overline{B}$ is associated with each of the sixteen places in each square (the same place in every square):

		(11)				(21)				(12)				(22)			
		(11)	(21)	(12)	(22)	(11)	(21)	(12)	(22)	(11)	(21)	(12)	(22)	(11)	(21)	(12)	(22)
(11)	(11)	$a'_{111}b'_{111}$	·	$a'_{111}b'_{121}$	·	$a'_{121}b'_{111}$	·	$a'_{121}b'_{121}$	·	·	·	·	·	·	·	·	·
	(21)	·	$a'_{112}b''_{111}$	·	$a'_{112}b''_{121}$	·	$a'_{122}b''_{111}$	·	$a'_{122}b''_{121}$	·	·	·	·	·	·	·	·
	(12)	$a'_{111}b'_{211}$	·	$a'_{111}b'_{221}$	·	$a'_{121}b'_{211}$	·	$a'_{121}b'_{221}$	·	·	·	·	·	·	·	·	·
	(22)	·	$a'_{112}b''_{211}$	·	$a'_{112}b''_{221}$	·	$a'_{122}b''_{211}$	·	$a'_{122}b''_{221}$	·	·	·	·	·	·	·	·
(21)	(11)	$a'_{211}b'_{111}$	·	$a'_{211}b'_{121}$	·	$a'_{221}b'_{111}$	·	$a'_{221}b'_{121}$	·	·	·	·	·	·	·	·	·
	(21)	·	$a'_{212}b''_{111}$	·	$a'_{212}b''_{121}$	·	$a'_{222}b''_{111}$	·	$a'_{222}b''_{121}$	·	·	·	·	·	·	·	·
	(12)	$a'_{211}b'_{211}$	·	$a'_{211}b'_{221}$	·	$a'_{221}b'_{211}$	·	$a'_{221}b'_{221}$	·	·	·	·	·	·	·	·	·
	(22)	·	$a'_{212}b''_{211}$	·	$a'_{212}b''_{221}$	·	$a'_{222}b''_{211}$	·	$a'_{222}b''_{221}$	·	·	·	·	·	·	·	·
(12)	(11)	·	·	·	·	·	·	·	·	$a''_{111}b'_{112}$	·	$a''_{111}b'_{122}$	·	$a''_{121}b'_{112}$	·	$a''_{121}b'_{122}$	·
	(21)	·	·	·	·	·	·	·	·	·	$a''_{112}b''_{112}$	·	$a''_{112}b''_{122}$	·	$a''_{122}b''_{112}$	·	$a''_{122}b''_{122}$
	(12)	·	·	·	·	·	·	·	·	$a''_{111}b'_{212}$	·	$a''_{111}b'_{222}$	·	$a''_{121}b'_{212}$	·	$a''_{121}b'_{222}$	·
	(22)	·	·	·	·	·	·	·	·	·	$a''_{112}b''_{212}$	·	$a''_{112}b''_{222}$	·	$a''_{122}b''_{212}$	·	$a''_{122}b''_{222}$
(22)	(11)	·	·	·	·	·	·	·	·	$a''_{211}b'_{112}$	·	$a''_{211}b'_{122}$	·	$a''_{221}b'_{112}$	·	$a''_{221}b'_{122}$	·
	(21)	·	·	·	·	·	·	·	·	·	$a''_{212}b''_{112}$	·	$a''_{212}b''_{122}$	·	$a''_{222}b''_{112}$	·	$a''_{222}b''_{122}$
	(12)	·	·	·	·	·	·	·	·	$a''_{211}b'_{212}$	·	$a''_{211}b'_{222}$	·	$a''_{221}b'_{212}$	·	$a''_{221}b'_{222}$	·
	(22)	·	·	·	·	·	·	·	·	·	$a''_{212}b''_{212}$	·	$a''_{212}b''_{222}$	·	$a''_{222}b''_{212}$	·	$a''_{222}b''_{222}$

We deal next with three sets of determinants:

$$
\begin{array}{l}
A^{(11)}, A^{(21)}, \ldots., A^{(l1)}, A^{(12)}, \ldots., A^{(l2)}, \ldots., A^{(1m)}, \ldots., A^{(lm)}, \\
B^{(11)}, B^{(21)}, \ldots., B^{(k1)}, B^{(12)}, \ldots., B^{(k2)}, \ldots., B^{(1m)}, \ldots., B^{(km)}, \\
C^{(11)}, C^{(21)}, \ldots., C^{(k1)}, C^{(12)}, \ldots., C^{(k2)}, \ldots., C^{(1l)}, \ldots., C^{(kl)};
\end{array}
$$

where

$$C^{(\alpha\beta)} \equiv \left| c^{(\alpha\beta)}_{\overset{\smile}{\gamma_1} \ldots. \overset{\smile}{\gamma_h} \overset{\frown}{\gamma_{h+1}} \ldots. \overset{\frown}{\gamma_r}} \right|^{(r)}_{m}$$

Forming now $U^{(1)} = A^{(11)} A^{(21)} \ldots. A^{(l1)} B^{(11)} B^{(21)} \ldots. B^{(k1)}$, and so on up to $U^{(m)} = A^{(1m)} \ldots. B^{(km)}$, construct $\overline{U} = U^{(1)} \ldots. U^{(m)}$ (as we did $\overline{A}$) so that

$$\overline{u}_{\overset{\smile}{(\alpha_1\beta_q\gamma)} \ldots. \overset{\smile}{(\alpha_f\beta_q\gamma)} \overset{\frown}{(\alpha_{f+1}\beta_q\gamma)} \ldots. \overset{\frown}{(\alpha_{p-1}\beta_q\gamma)} \overset{\smile}{(\alpha_p\beta_1\gamma)} \ldots. \overset{\smile}{(\alpha_p\beta_g\gamma)} \overset{\frown}{(\alpha_p\beta_{g+1}\gamma)} \ldots. \overset{\frown}{(\alpha_p\beta_{q-1}\gamma)}} \equiv a^{(\beta_q\gamma)}_{\alpha_1 \ldots. \alpha_p} \, b^{(\alpha_p\gamma)}_{\beta_1 \ldots. \beta_q};$$

and construct $\overline{C}$ (as we did $\overline{B}$) so that

$$\overline{c}_{\overset{\smile}{(\alpha\beta\gamma_1)} \ldots. \overset{\smile}{(\alpha\beta\gamma_h)} \overset{\frown}{(\alpha\beta\gamma_{h+1})} \ldots. \overset{\frown}{(\alpha\beta\gamma_r)}} \equiv c^{(\alpha\beta)}_{\gamma_1 \ldots. \gamma_r}.$$

Multiply together $\overline{U}$ and $\overline{C}$ by rows into a determinant V of class $p+q+r-4$ and order klm. The elements of V which are not zeros are monomials, and $\alpha_p = \alpha$, $\beta_{q-1} = \beta$, $\gamma_r = \gamma$. Detaching the suffixes of α, β and γ in the locant of an element of V, we have the prescription:

$$v \left\{ \begin{array}{llllllll} \alpha: & \overset{\smile}{1} \ldots \overset{\smile}{f} & \overset{\frown}{f+1} \ldots \overset{\frown}{p-1} & \overset{\smile}{p} \ldots \overset{\smile}{p} & \overset{\frown}{p} \ldots \overset{\frown}{p} & \overset{\smile}{p} \ldots \overset{\smile}{p} & \overset{\frown}{p} \ldots \overset{\frown}{p} \\ \beta: & q \ldots q & q \ldots q & 1 \ldots g & g+1 \ldots q-2 & q-1 \ldots q-1 & q-1 \ldots q-1 \\ \gamma: & r \ldots r & r \ldots r & r \ldots r & r \ldots r & 1 \ldots h & h+1 \ldots r-1 \end{array} \right\}$$

$$\equiv a^{(\beta_q\gamma_r)}_{\alpha_1 \ldots. \alpha_p} \, b^{(\gamma_r\alpha_p)}_{\beta_1 \ldots. \beta_q} \, c^{(\alpha_p\beta_{q-1})}_{\gamma_1 \ldots. \gamma_r}.$$

Here at least the indices $(\alpha_{p-1}\beta_q\gamma_r)$ and $(\alpha_p\beta_{q-1}\gamma_{r-1})$ are signant. The superfixes on the right are the sets of two consecutive indices in the sequence

$$\beta_q \, \gamma_r \, \alpha_p \, \beta_{q-1}.$$

The prescription applicable to *four* sets of determinants is obtainable by: (i) changing v to w; (ii) subjoining to the locant the line

$$\delta: \; s \ldots. s \ldots. s \ldots. s \; 1 \ldots. i \;\; i+1 \ldots. s-1,$$

the last s falling under γ_{r-2}, the 1 under γ_{r-1}, and the values p, $q-1$, and $r-1$ being continued over $\delta_2 \ldots. \delta_{s-1}$; (iii) annexing on the right

$$d^{(\alpha_p\beta_{q-1}\gamma_{r-1})}_{\delta_1 \ldots. \delta_s};$$

and (iv) inserting δ_s after γ_r and before α_p in the first three superfixes, so that the superfixes are now the sets of three consecutive indices in the sequence

$$\beta_q \, \gamma_r \, \delta_s \, \alpha_p \, \beta_{q-1} \, \gamma_{r-1}.$$

Each prescription is obtained in like manner from the prescription that precedes, and we have the following general theorem:

The product of r sets of determinants

$$A^{(hg_h)} \equiv \left| a^{(hg_h)}{}_{\breve{\alpha}_{h1} \breve{\alpha}_{h2} \dots \breve{\alpha}_{hf_h} \hat{\alpha}_{hf_h+1} \dots \hat{\alpha}_{hp_h}} \right|_{n_h}^{(p_h)},$$

where $h=1, 2, \dots, r$, *and where* g_h *is the h-th set of* $r-1$ *consecutive indices in the sequence*

$$\alpha_{2p_2}\alpha_{3p_3} \dots \alpha_{rp_r}\alpha_{1p_1}\alpha_{2,\, p_2-1}\alpha_{3,\, p_3-1} \dots \alpha_{r-1,\, p_{r-1}-1},$$

is expressible as a determinant A *of class* $p_1+\dots+p_r-2(r-1)$ *and order* $n_1 n_2 \dots n_r$, *in which all elements are zeros excepting those given by the prescription*

$$a \left\{ \begin{matrix} \breve{(\alpha_{11} g_1)} \dots \breve{(\alpha_{1f_1} g_1)}\ \widehat{(\alpha_{1f_1+1} g_1)} \dots \widehat{(\alpha_{1p_1-1} g_1)} \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \breve{(\alpha_{k1} g_k)} \dots \breve{(\alpha_{kf_k} g_k)}\ \widehat{(\alpha_{kf_k+1} g_k)} \dots \widehat{(\alpha_{kp_k-2} g_k)} \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \breve{(\alpha_{r1} g_r)} \dots \breve{(\alpha_{rf_r} g_r)}\ \widehat{(\alpha_{rf_r+1} g_r)} \dots \widehat{(\alpha_{rp_r-1} g_r)} \end{matrix} \right\} \equiv \prod_{h=1}^{r} a^{(hg_h)}_{\alpha_{h1} \dots \alpha_{hp_h}},$$

where $k=2, 3, \dots, r-1$.

Remark that at least the indices $(\alpha_{1p_1-1} g_1)$ and $(\alpha_{rp_r-1} g_r)$ are signant.

If all the determinants are of two dimensions, the prescription takes the form

$$a_{(\alpha_{11}\alpha_{22}\alpha_{32} \dots \alpha_{r2})(\alpha_{12}\alpha_{21}\alpha_{31} \dots \alpha_{r-1,1}\alpha_{r1})} \equiv \prod_{h=1}^{r} a^{(hg_h)}_{\alpha_{h1}\alpha_{h2}},$$

the sequence to which g_h applies being

$$\alpha_{22}\alpha_{32} \dots \alpha_{r2}\alpha_{12}\alpha_{21}\alpha_{31} \dots \alpha_{r-1,1}.$$

This agrees with the theorem of Metzler designated T_s by Moore.

12. *An Application to Transvectants.**

THEOREM: *If the* n^{p+q} *k-th transvectants of all possible pairs of binary forms taken from two sets*

$$\| f_{\alpha_1 \dots \alpha_p} \|_n^{(p)}, \quad \| \phi_{\beta_1 \dots \beta_q} \|_n^{(q)}, \quad n \geqq k+2,$$

be made the elements of a determinant

$$C \equiv \left| c_{\breve{\alpha}_1 \dots \breve{\alpha}_g \hat{\alpha}_{g+1} \dots \hat{\alpha}_p \breve{\beta}_1 \dots \breve{\beta}_h \hat{\beta}_{h+1} \dots \hat{\beta}_q} \right|_n^{(p+q)}, \quad \begin{cases} p-g \text{ odd}, \\ q-h \text{ odd}, \end{cases}$$

with

$$c_{\alpha_1 \dots \alpha_p \beta_1 \dots \beta_q} \equiv (f_{\alpha_1 \dots \alpha_r}, \phi_{\beta_1 \dots \beta_q})^k,$$

then

$$C \equiv 0.$$

---

* Special cases of this theorem have been given by L. Gegenbauer (*loc. cit.*); see Abrégé, p. 94, where an important restriction on Gegenbauer's results, not mentioned by him, is brought out.

PROOF: We can form two null determinants A and B, whose elements are k-th derivatives of the f's and ϕ's respectively, with suitable numerical factors, such that $A \cdot B = C$. Letting $m_{a_1 \dots a_p}$, $\mu_{\beta_1 \dots \beta_q}$ be the degrees of $f_{a_1 \dots a_p}$, $\phi_{\beta_1 \dots \beta_q}$, respectively, set up:

$$A \equiv | a_{\breve{a}_1 \dots \breve{a}_g \hat{a}_{g+1} \dots \hat{a}_p \hat{\lambda}} |_n^{(p+1)}, \quad a_{a_1 \dots a_p \lambda} \equiv \frac{(m_{a_1 \dots a_p} - k)!}{m_{a_1 \dots a_p}!} \cdot \frac{\partial^k f_{a_1 \dots a_p}}{\partial_{x_1}^{k-\lambda+1} \partial_{x_2}^{\lambda-1}};$$

$$B \equiv | b_{\breve{\beta}_1 \dots \breve{\beta}_h \hat{\beta}_{h+1} \dots \hat{\beta}_q \hat{\lambda}} |_n^{(q+1)}, \quad b_{\beta_1 \dots \beta_q \lambda} \equiv (-1)^{\lambda-1} \binom{k}{\lambda-1} \frac{(\mu_{\beta_1 \dots \beta_q} - k)!}{\mu_{\beta_1 \dots \beta_q}!} \cdot \frac{\partial^k \phi_{\beta_1 \dots \beta_q}}{\partial_{x_1}^{\lambda-1} \partial_{x_2}^{k-\lambda+1}}.$$

These prescriptions fill only the first $k+1$ places in each row (file of the last direction), and we shall fill the remaining places in each row with zeros, the result being that one or more layers in A and in B will consist of zeros, whence $A=0$ and $B=0$. Obviously $A \cdot B = C$, the multiplication being of row into row.

COROLLARY 1. *If the n^4 k-th transvectants of all possible pairs of binary forms taken from two sets*

$$\left\| \begin{matrix} f_{11} \cdots f_{1n} \\ \cdots\cdots \\ f_{n1} \cdots f_{nn} \end{matrix} \right\|, \quad \left\| \begin{matrix} \phi_{11} \cdots \phi_{1n} \\ \cdots\cdots \\ \phi_{n1} \cdots \phi_{nn} \end{matrix} \right\|, \quad n \geqq k+2.$$

be made the elements of a 4-way determinant with two signant indices:

$$C \equiv | c_{\breve{a}_1 \hat{a}_2 \breve{\beta}_1 \hat{\beta}_2} |_n^{(4)}, \quad c_{a_1 a_2 \beta_1 \beta_2} \equiv (f_{a_1 a_2}, \phi_{\beta_1 \beta_2})^k,$$

then $$C \equiv 0.$$

COROLLARY 2. *If the n^2 k-th transvectants of all possible pairs of binary forms taken from two sets*

$$f_1, f_2, \dots, f_n, \quad \phi_1, \phi_2, \dots, \phi_n, \quad n \geqq k+2,$$

be made the elements of a 2-way determinant

$$C \equiv \left| \begin{matrix} (f_1, \phi_1)^k \cdots (f_1, \phi_n)^k \\ \cdots\cdots\cdots \\ (f_n, \phi_1)^k \cdots (f_n, \phi_n)^k \end{matrix} \right|_n,$$

then $$C \equiv 0.$$

This corollary includes as a special case Gordan's result:

$$\left| \begin{matrix} (f_1, \phi_1)^2 \cdots (f_1, \phi_4)^2 \\ \cdots\cdots\cdots \\ (f_4, \phi_1)^2 \cdots (f_4, \phi_4)^2 \end{matrix} \right|_4 \equiv 0.$$

CAMBRIDGE, MASSACHUSETTS.

*On a Certain General Class of Functional Equations.**

By W. Harold Wilson.

---

§ 1. *Introduction and General Considerations.*

Addition formulae of the general type

$$G[f(x+y), f(x), f(y)] = 0,$$

where G is a polynomial in its three arguments, play a prominent rôle in the theory of elliptic functions. A natural generalization of such formulae is

$$P[x, y, f(x), f(y), f(\alpha_1 x+\beta_1 y), \ldots, f(\alpha_n x+\beta_n y)] = 0, \qquad \text{(I)}$$

where

(i) P denotes a polynomial in its $n+4$ arguments such that every argument involving f is explicitly present;

(ii) x and y are independent variables;

(iii) α_i and β_i, $i=1, 2, \ldots, n$, are given constants;† and,

(iv) $f(x)$ is an unknown single-valued function to be determined so that equation (I) shall be identically satisfied.‡

Equation (I) *is said to be of order n. The degree m of P in the function f is said to be the degree of equation* (I).

---

* Read before the American Mathematical Society (at Chicago), April 6, 1917.

† For the purposes of this paper it is convenient to carry certain hypotheses in regard to the α's and β's. A statement of these hypotheses is to be found below.

‡ A theorem of some interest in the general theory of these functional equations is that *every solution $f(x)$ of equation* (I) *is a solution of a similar equation in which x and y occur only in the arguments of the function f.* To prove this arrange P as a polynomial in x and y. The substitutions $x = s + k_i t$, $y_i = t$, where $k_0 = 0$ and

$$k_i \neq k_\lambda \pm \frac{\beta_h}{\alpha_h}, \quad k_i \neq k_\lambda + \frac{\beta_h}{\alpha_h} - \frac{\beta_j}{\alpha_j},$$

$\lambda = 0, 1, \ldots, i-1, h, j = 1, 2, \ldots, n$, transform (I) into equations similar to (I) such that the highest degree in s and t is the same for all of them. It is easily seen that a finite number of non-zero k's may be employed such that the variables s and t may be eliminated from these equations, in so far as they occur as coefficients, by Sylvester's dialytic method of elimination. The result of this elimination is an equation (II) which states that a polynomial Q in the function f has the value zero. The arguments of f are linear combinations of two independent variables s and t. Hence (II) is similar to (I), although in general its order and degree will differ from those of (I). Furthermore, if no two arguments of f in (I) are proportional, then no two arguments of f in (II) are proportional.

Cauchy* discussed two special cases of (I) and two related equations, namely:

$$f(x+y)=f(x)+f(y),\quad f(x+y)=f(x)f(y),$$
$$f(xy)=f(x)+f(y),\quad f(xy)=f(x)f(y).$$

One or the other of the first two of these equations has since been treated † by Darboux, E. B. Wilson, Vallée Poussin, Schimmack and Hamel. Carmichael ‡ has given a generalization of the Cauchy equations while Jensen § has discussed several applications of them. Cauchy || has treated the equation

$$\phi(x+y)+\phi(x-y)=2\phi(x)\phi(y).$$

Carmichael ¶ has considered the equations

$$h(x+y)h(x-y)=h^2(x)+h^2(y)-c^2,\quad g(x+y)g(x-y)=g^2(x)-g^2(y).$$

Van Vleck and H'Doubler ** have discussed the equation

$$\psi(x+y)\psi(x-y)=[\psi(x)\psi(y)]^2.$$

Other related equations have been considered by several writers, and systems of functional equations have also been treated.

It seems that no systematic account of a general theory for equations of the form (I) has ever been undertaken. This paper is designed to contribute to such an account. The equations considered in the principal part (§§ 2 to 8) of the paper are linear homogeneous equations with constant coefficients. They may be written in the form

$$\sum_{i=1}^{n}\gamma_i f(\alpha_i x+\beta_i y)+\gamma_{n+1}f(x)+\gamma_{n+2}f(y)=0. \tag{1}$$

It will be shown that if some α's and β's having different subscripts are zero and no ratio α_i/β_i of non-zero α's and β's is distinct from all the remaining ratios, the equation is exceptional. The exceptional case receives mention only in §§ 6 and 11. There is no loss of generality in assuming that no α is zero in the non-exceptional case. For convenience in exposition the hypothesis will be carried in the text that in addition to no α being zero, no β is zero, and no two ratios α_i/β_i are equal. The additional argumentation for the remaining non-exceptional cases is supplied in footnotes.

---

* *Cours d'Analyse* (1821), Chapter 5. Cauchy treated the last two equations by transforming them into the first two equations. It is obvious that similar transformations may be applied to reduce more general equations to the form of those considered in this paper.

† Darboux, *Mathematische Annalen*, Vol. XVII (1880), p. 56. E. B. Wilson, *Annals of Mathematics*, Vol. I, Ser. 2 (1899), p. 47. Vallée Poussin, *Cours d'Analyse infinitésimale* (1903), p. 30. Schimmack, *Nova Acta*, Vol. XC, p. 5. Hamel, *Mathematische Annalen*, Vol. LX (1905), p. 459.

‡ *American Mathematical Monthly*, Vol. XVIII (1911), p. 198.

§ *Tidsskrift for Mathematik*, Vol. II, Ser. 4 (1878), p. 149.

|| *Cours d'Analyse* (1821), p. 114.

¶ *American Mathematical Monthly*, Vol. XVI (1909), p. 180.

** *Transactions of the American Mathematical Society*, Vol. XVII (1916), p. 9.

A normal equation of order n is derived (in § 2) which is satisfied by every solution of any non-exceptional equation (1) of order n. This normal equation forms a foundation upon which the entire development of the theory of equation (1) is based. Any normal solution may be uniquely determined at each point of a dense set covering the complex plane if it is given at the vertices of a certain triangular network (§§ 3, 4). It is shown in § 5 that *the normal solution analytic in the neighborhood of the point zero of the complex plane is an arbitrary polynomial in x of degree n.* It is also shown that *the normal solution continuous in the neighborhood of the point zero of the complex plane is an arbitrary polynomial in u and v of degree n where u and v are real and $x=u+v\sqrt{-1}$.* The normal solution analytic along any line in the finite complex plane is also an arbitrary polynomial in x of degree n and the normal solution continuous along any line in the finite complex plane is an arbitrary polynomial in u of degree n if the line is not parallel to the axis of imaginaries and an arbitrary polynomial in v of degree n if the line is not parallel to the axis of reals. The analytic and continuous solutions of (1) are found (§ 6) from the normal solution. Examples are exhibited in § 6 which show that equations of type (1) may have non-trivial continuous solutions, but no non-trivial analytic solutions, while other examples show that equations of type (1) may have analytic solutions which are also the most general continuous solutions. A converse theorem is briefly considered in § 7. It is shown (in § 8) that *if a function $f(x)$ satisfying an equation of type* (1) *has a point of discontinuity in the finite complex plane* [*or on any line in the finite complex plane*] *it has a point of discontinuity in every region* [*interval*] *of the plane* [*line*], *however small.*

Equation (1) is employed (§ 9) to solve certain equations of the type

$$\sum_{i=1}^{n}\phi_i(x, y)f(\alpha_i x+\beta_i y)+\phi_{n+1}(x, y)f(x)+\phi_{n+2}(x, y)f(y)+\phi_{n+3}(x, y)=0,$$

where the ϕ's are known functions. Equation (1) is also employed (§ 10) to find all analytic solutions, and in some cases, the continuous solutions of binomial equations of the type

$$\prod_{i=1}^{k}[f(\alpha_i x+\beta_i y)]^{\gamma_i}=C\prod_{i=k+1}^{n+1}[f(\alpha_i x+\beta_i y)]^{\gamma_i}[f(y)]^{\gamma_{n+2}},$$

where no α is zero, C is a constant and the γ's are constants of which the real parts are positive. Pexider* used the first Cauchy equation to solve

$$f(x)+\phi(y)=\psi(x+y).$$

In § 11 the method of obtaining the solutions of (1) is used to solve the equation

$$\sum_{i=1}^{n}\gamma_i f_i(\alpha_i x+\beta_i y)+\gamma_{n+1}f_{n+1}(x)+\gamma_{n+2}f_{n+2}(y)=0,$$

* *Monatshefte für Mathematik und Physik*, Vol. XIV (1903), p. 293.

of which the equation considered by Pexider is a special case. It is proved that when no two arguments of f in the foregoing equation are proportional, each continuous solution f is a polynomial of degree not greater than n.

§ 2. *Reduction to a Normal Form.*

The solution $f(x)$ of the general n-th order equation

$$\sum_{i=1}^{n}\gamma_i f(\alpha_i x+\beta_i y)+\gamma_{n+1}f(x)+\gamma_{n+2}f(y)=0, \tag{1}$$

is contained in that of an equation of the same form and same order [equation (6) below] in which each α, β, γ is a given integer. The derivation of (6) from (1) is accomplished by elimination.

If (1) is subtracted from the equation derived from (1) by replacing y by $y+t_{n+1}$, the result is

$$\sum_{i=1}^{n}\gamma_i[f(\alpha_i x+\beta_i y+\beta_i t_{n+1})-f(\alpha_i x+\beta_i y)]+\gamma_{n+2}[f(y+t_{n+1})-f(y)]=0. \tag{2}$$

If (2) is subtracted from the equation derived from (2) by replacing x by $x-\beta_1 t_1$ and y by $y+\alpha_1 t_1$, the result is

$$\begin{aligned}\sum_{i=2}^{n}\gamma_i[&f(\alpha_i x+\beta_i y+\Delta_{1i}t_1+\beta_i t_{n+1})-f(\alpha_i x+\beta_i y+\Delta_{1i}t_1)\\ &-f(\alpha_i x+\beta_i y+\beta_i t_{n+1})+f(\alpha_i x+\beta_i y)]\\ &+\gamma_{n+2}[f(y+\alpha_1 t_1+t_{n+1})-f(y+\alpha_1 t_1)-f(y+t_{n+1})+f(y)]=0,\end{aligned} \tag{3}$$

where

$$\Delta_{ji}=\alpha_j\beta_i-\alpha_i\beta_j.$$

It is easily seen that this is true because the given substitutions leave $\alpha_1 x+\beta_1 y$ unchanged, but replace $\alpha_i x+\beta_i y$, $i\neq 1$, by

$$\alpha_i x+\beta_i y+\alpha_1\beta_i t_1-\alpha_i\beta_1 t_1=\alpha_i x+\beta_i y+\Delta_{1i}t_1.$$

In general, if an equation (a) results after such eliminations, then each argument of f in (a) is a linear expression in x, y and certain t's, the subscripts of the t's corresponding to those in the terms eliminated. The substitution of

$$x-\beta_j t_j \text{ for } x \text{ and } y+\alpha_j t_j \text{ for } y \tag{4}$$

gives rise to an equation (b) which differs from (a) by having each $\alpha_i x+\beta_i y$ of (a) replaced by $\alpha_i x+\beta_i y+\Delta_{ji}t_j$. Since $\Delta_{jj}=0$ and since (4) does not affect the t's that are found in (a), it follows that the difference formed by subtracting (a) from (b) contains no term for which i is equal to the fixed integer j. Since n is finite these eliminations may be continued until only terms having the coefficients $\pm\gamma_{n+2}$ remain. Moreover, the order of elimination of the terms for which $i=1, 2, \ldots, n$, is immaterial.

It is obvious that the equation resulting from these eliminations is linear and homogeneous. Since $\gamma_{n+2} \neq 0$ by the assumption that (1) is of order n, the equation may be simplified by dividing by γ_{n+2}. It is easily seen that there are $(n+1)!/k!(n+1-k)!$ distinct terms in which the coefficient of f is $(-1)^k$, and in which the argument of f is obtained by omitting k terms after the first from $y+\alpha_1 t_1+\alpha_2 t_2+\ldots\ldots+\alpha_n t_n+t_{n+1}$. Moreover, no other such terms are possible. This is true for $k=0, 1, \ldots\ldots, n+1$. The substitutions t_1 for $\alpha_1 t_1$, t_2 for $\alpha_2 t_2, \ldots\ldots, t_n$ for $\alpha_n t_n$, t_{n+2} for y, serve to completely determine an equation independent of the original α's, β's and γ's. If Σ_k denotes the sum of the $(n+1)!/k!(n+1-k)!$ terms in which the arguments are formed by omitting k t's from $\Sigma_{i=1}^{n+1} t_i$, then the equation which is satisfied by $f(x)$ is

$$\Sigma_0-\Sigma_1+\Sigma_2-\ldots\ldots+(-1)^n\Sigma_n+(-1)^{n+1}\Sigma_{n+1}=0. \tag{5}$$

Equation (5) involves $n+2$ independent variables. In order to obtain a normalized equation having the same form and order as (1), let $t_1=t_2=\ldots\ldots=t_{n+1}=x$ and $t_{n+2}=y$, whence (5) becomes*

$$f[(n+1)x+y]+\ldots\ldots+(-1)^{n+1-k}\frac{(n+1)!}{k!(n+1-k)!}f(kx+y)$$
$$+\ldots\ldots+(-1)^n f(x+y)+(-1)^{n+1}f(y)=0. \tag{6}$$

From the foregoing considerations we see that *every solution $f(x)$ of equation* (1) *is a solution of equation* (5) *and of the normal equation* (6).

While the solutions $f(x)$ of (1) are included among those of (6), it is not necessarily true that the solutions of (6) are included in those of (1). The following example suffices to prove this statement. Equation (6) for $n=2$ is

$$f(3x+y)-3f(2x+y)+3f(x+y)-f(y)=0. \tag{7}$$

As will be shown in § 5, the most general continuous solution of (7) over the finite complex x-plane is

$$f(x)=au^2+buv+cv^2+du+ev+f, \tag{8}$$

where $x=u+v\sqrt{-1}$ (u and v real) and a, b, c, d, e and f are arbitrary constants. From the above considerations we see that the solution of any second order equation of form (1) is included among those of (8). However, substitution shows that for the equation

$$f(2x+y)-2f(x+y)-2f(x)+f(y)=0$$

---

* The results in (5) and (6) can be rendered more precise in special cases. If $\beta_j/a_j=\beta_h/a_h$, $\Delta_{jh}=0$ and the substitution of $x-\beta_j t_j$ for x and $y+a_j t_j$ for y leaves $a_h x+\beta_h y$, as well as $a_j x+\beta_j y$, unchanged. Hence the elimination of terms for which $i=j$ also eliminates all terms for which $i=h$ where h is any value for which $\Delta_{jh}=0$. But when $\Delta_{jh}=0$, $\beta_j/a_j=\beta_h/a_h$, and therefore this proportionality is a necessary and sufficient condition for the simultaneous elimination of terms for more than one value of i. Hence *the solution of any m-th order equation* (1) *in which no a is zero, is contained in the solution of equation* (5) *or of equation* (6) *where n is the number of distinct ratios β_i/a_i in* (1).

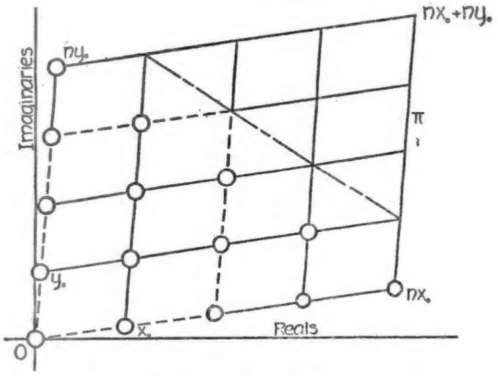
we have $d=e=f=0$. In fact, such an equation need have no other than the trivial solution $f(x)\equiv 0$ as is shown by the equation

$$f(2x+y)-f(x+y)-2f(x)+f(y)=0.$$

§ 3. *Determination of the Normal Solution at the Vertices of a Network.*

The function $f(x)$ to be considered in this section and in §§ 4, 5 is the solution of the normal equation developed in § 2. It is to be observed that if f is given for the arguments ky_0, x_0+ky_0, $2x_0+ky_0$, $\ldots$, nx_0+ky_0, then f is known for the argument $(n+1)x_0+ky_0$ by (6). By putting $y=jx_0+ky_0$ for successive values $1, 2, \ldots,$ of j, f is determined by (6) for any argument mx_0+ky_0, where m is any positive integer. Similarly, by successively giving to j the values $-1, -2, \ldots,$ f is determined by (6) for any argument mx_0+ky_0, where m is a negative integer. By interchanging x and y in (6) it is easily seen that f is known for all arguments mx_0+ky_0, where k is any integer or zero, if it is given for the arguments mx_0, mx_0+y_0, mx_0+2y_0, $\ldots$, mx_0+ny_0. Hence, if f is given for the arguments mx_0+ky_0, $m, k=0, 1, \ldots, n$, then it is known by (6) for all arguments mx_0+ky_0, where m is any integer whatever and $k=0, 1, \ldots, n$, and finally, it is known by (6) for all arguments mx_0+ky_0, where m and k are any integers whatever.

It will now be proved that if f is known for the arguments mx_0+ky_0, where m and k are zero or positive integers and $m+k<n+1$, then it is also known for the arguments mx_0+ky_0, where $m, k=0, 1, \ldots, n$. The figure illustrates the case for which $n=4$. The points designated by the small circles



represent the arguments for which f is supposed known. The parallelogram formed by the lines joining the points 0, nx_0, nx_0+ny_0, and ny_0 will be denoted by π. The parallelogram π contains, on and within its boundary, all the points mx_0+ky_0, $m, k=0, 1, \ldots, n$, and no other points as vertices. It is required, therefore, to find f at the vertices of π not designated by small circles.

Equation (5) furnishes a simple means of solution of the present problem. Let $t_1=t_2=\ldots\ldots=t_q=x_0$, $t_{q+1}=\ldots\ldots=t_{n+1}=y_0$ and $t_{n+2}=0$. By these substitutions every argument in (5) is of the form mx_0+ky_0 where m and k are zero or positive integers and $m+k<n+1$, except in the first term in which $m+k=n+1$. Moreover, $m\leqq q$ and $k\leqq n+1-q$; hence f is determined for the argument $qx_0+(n+1-q)y_0$ in terms of linear combinations of its values for arguments represented by vertices within or on the boundary of the parallelogram bounded by lines joining the four points 0, qx_0, $qx_0+(n+1-q)y_0$, $(n+1-q)y_0$, (represented in the figure by dotted lines for $q=2$). By giving q the values 1, 2, ..., n, in succession, f is determined at each vertex of π which lies on the line joining the points x_0+ny_0 and nx_0+y_0 (the dot-and-dash line in the figure). It is to be noticed that each determination is made from a linear equation in one unknown with unit coefficient. Hence, the functional values so found are unique.

Having determined f for the arguments mx_0+ky_0 in π for which $m+k=n+1$, it is easy to determine f at the vertices for which $m+k=n+2$. In (5) let $t_1=t_2=\ldots\ldots=t_q=x_0$ and $t_{q+1}=\ldots\ldots=t_{n+2}=y_0$. Thus f is determined for the argument $qx_0+(n+2-q)y_0$ in terms of arguments represented by vertices within and on the boundary of the parallelogram formed by the four lines joining the points 0, qx_0, $qx_0+(n+2-q)y_0$, $(n+2-q)y_0$. This determination is also made by means of a linear equation involving but one unknown with the coefficient $+1$. Hence, no indetermination can be introduced. By giving q the successive values 2, 3, ..., n, f is determined for all arguments in π which lie on the line joining the points $2x_0+ny_0$ and nx_0+2y_0.

Proceeding in this manner f is determined for all the remaining arguments of π by giving t_{n+2} the values $2y_0$, $3y_0$, ..., $(n-1)y_0$ successively. For each value hy_0 of t_{n+2}, f must be determined for all arguments

$$qx_0+(n+h+1-q)y_0,\quad q=h+1,\quad h+2,\ \ldots,\ n,$$

before giving t_{n+2} the value $(h+1)y_0$. As before, each determination is made by a linear equation in one unknown with coefficient $+1$. Hence, f is uniquely determined at all the vertices of π, and we have the following result:

Every solution $f(x)$ of the normal equation (6) *is known for the points mx_0+ky_0 where m and k are any integers of zero if it is given at the points mx_0+ky_0 for which m and k are positive integers or zero and $m+k<n+1$.*

§ 4. *Determination of the Normal Solution at a Dense Set of Points.*

It will now be shown that if the solution of the normal equation is known at all points mx_0+ky_0, where m and k are integers, then it may be found for

$\frac{1}{2}mx_0+\frac{1}{2}ky_0$, and finally for $2^{-s}mx_0+2^{-s}ky_0$ for all integers m, k and s. If f is found for certain appropriate linear combinations of $2^{-s}x_0$ and $2^{-s}y_0$, then it is evident that $2^{-s}x_0$ and $2^{-s}y_0$ may be regarded as were x_0 and y_0 in § 3. Hence, if it is most convenient to determine f at $2^{-s}mx_0+2^{-s}ky_0$, $m, k=0, 1, \ldots, 2n$, it is also sufficient, in view of the argument of § 3, for f will then be known at these points for all integers m and k. Furthermore, if it is shown that f is known at $2^{-1}mx_0+2^{-1}ky_0$ for all integers m and k when it is known at mx_0+ky_0, then by regarding $2^{1-s}x_0$ and $2^{1-s}y_0$ as x_0 and y_0 for the successive values $s=2, 3, \ldots$, it is seen that f is known at $2^{-s}mx_0+2^{-s}ky_0$ for all integers m, k and s. Therefore, since the points $2^{-s}mx_0+2^{-s}ky_0$, m, k and s any integers whatever, form a dense set, it is sufficient to show that f is known at $\frac{1}{2}mx_0+\frac{1}{2}ky_0$, $m, k=0, 1, \ldots, 2n$, to show that f is known at all points of a dense set.

Supposing f known at $\frac{1}{2}mx_0+\frac{1}{2}ky_0$, $m, k=0, 2, \ldots, 2n$, it is only required to learn f at $\frac{1}{2}mx_0+\frac{1}{2}ky_0$, $m, k=1, 3, \ldots, 2n-1$. Replacing x by $\frac{1}{2}x_0$ and y by $\frac{1}{2}hx_0+ky_0$ in (6), n equations are obtained by giving h the values $n-1, \ldots, 1, 0$. The equations are linear in f. Alternate terms involve the arguments $\frac{1}{2}mx_0+ky_0$, $m=1, 3, \ldots, 2n-1$, for which f is unknown. The n equations may be considered as linear equations in n unknowns. It is easily seen that the determinant of the coefficients of the unknowns is $(-1)^{\nu}\Delta_n$, where ν is $n/2$ or $(n+1)/2$ according as n is even or odd, and Δ_n is a determinant of binomial coefficients such that the element in the i-th row and j-th column is

$$(n+1)!/(2j-i)!(n-2j+i+1)!,$$

unless $2j<i$ or $2j>n+i+1$, in which case it is zero. Subtracting the $(i+1)$-th row from the i-th row, i having the values $1, 2, \ldots, n-1$, in order, the element in the i-th row and j-th column, $i\neq n$, becomes

$$(n-4j+2i+2)\ (n+1)!/(2j-i)!(n-2j+i+2)!,$$

unless $2j<i$ or $2j>n+i+2$, in which case it is zero. Adding to the j-th column the sum of the preceding columns, giving j the values $n, n-1, \ldots, 2$ in order, Δ_n assumes a form in which the element in the i-th row and j-th column, $i\neq n$, is

$$S=\sum_{h=1}^{j}(n-4h+2i+2)(n+1)!/(2h-i)!(n-2h+i+2)!.$$

If for any given value of j this sum is $n!/(2j-i)!(n-2j+i)!$, unless $2j<i$ or $2j>n+i$, in which case it is zero, it is easily shown that for $j+1$ it is $n!/(2j-i+2)!(n-2j+i-2)!$, unless $2j<i-2$ or $2j>n+i-2$, in which case it is zero. For $2j-i=0$ or $2j-i=1$ the value of S is merely the first non-vanishing term, that is, 1 or n, respectively. Since $n!/(2j-i)!(n-2j+i)!$ is

equal to 1 for $2j-i=0$, and equal to n for $2j-i=1$, it follows by induction that S has the value

$$S=(n)!/(2j-i)!(n-2j+i)!,$$

unless $2j<i$ or $2j>n+i$, in which case it is zero. For the n-th column the value of S is always zero for $i\neq n$ since $2j>n+i$. For $i=n$, it is clear that the elements are affected only by the process of addition, and since every alternate term of $(1+1)^{n+1}$ is involved, the last term in the n-th column is 2^n. Hence, Δ_n has been so transformed that the principal $(n-1)$-rowed minor found by deleting the last row and last column of Δ_n is Δ_{n-1}, and the last column consists exclusively of zeros with the single exception of the element 2^n in the n-th row. Therefore $\Delta_n=2^n\Delta_{n-1}$. It is evident that $\Delta_2=2^{1+2}$ (Δ_1 is trivially 2). Therefore $\Delta_n=2^{1+2+\dots+n}=2^{\frac{1}{2}n(n+1)}$ which is distinct from zero for all finite values of n. Since $(-1)^\nu\Delta_n\neq 0$, it follows immediately that f is uniquely determined at the points $\frac{1}{2}mx_0+ky_0$, $m=1, 3, \dots, 2n-1$, and hence for $m=1, 2, \dots, 2n$.

Having determined f at the points $\frac{1}{2}mx_0+ky_0$, $m=1, 2, \dots, 2n$, it is only necessary to let $x=\frac{1}{2}y_0$ and $y=\frac{1}{2}hy_0+\frac{1}{2}mx_0$ in (6), letting h take the values $n-1, \dots, 1, 0$, to determine f at the points $\frac{1}{2}mx_0+\frac{1}{2}ky_0$, $m, k=0, 1, \dots, 2n$, for $(-1)^\nu\Delta_n$ is again the determinant of the coefficients of the unknown terms.

Combining the results of §§ 3 and 4, it may be stated that f *is determined at each point of the dense set* $2^{-s}mx_0+2^{-s}ky_0$ *if it is known at the points* mx_0+ky_0 *for which* m *and* k *are positive integers or zero and* $m+k<n+1$.

§ 5. *Solutions of the Normal Equation.*

Equation (6) may be written in the form

$$\sum_{j=0}^{n+1}(-1)^j\frac{(n+1)!}{j!(n-j+1)!}f[(ju+s)+\sqrt{-1}(jv+t)]=0,$$

where $x=u+v\sqrt{-1}$, $y=s+t\sqrt{-1}$, and u, s, v, t are real. Let us seek a solution $f(x)$ of this equation in the form

$$f(x)=\sum_{h=0}^{N}\sum_{q=0}^{h}c_{qh}u^qv^{h-q}, \tag{9}$$

in which the coefficients c_{qh} are constants. Putting this value of $f(x)$ in (6) we have

$$\sum_{j=0}^{n+1}(-1)^j\frac{(n+1)!}{j!(n-j+1)!}\sum_{h=0}^{N}\sum_{q=0}^{h}c_{qh}(ju+s)^q(jv+t)^{h-q}=0.$$

The terms involving $u^{q-p}s^pv^{h-q-r}t^r$ are

$$\sum_{j=0}^{n+1}\left[(-1)^j\frac{(n+1)!}{j!(n-j+1)!}j^{h-p-r}\right]c_{qh}\frac{q!}{p!(q-p)!}\frac{(h-q)!}{r!(h-q-r)!}u^{q-p}s^pv^{h-q-r}t^r.$$

Since no term outside of the brackets involves j, it is evident that the given expression vanishes, for non-zero c_{qh}, u, s, v and t, when and only when

$$B_{qh} \equiv \sum_{j=0}^{n+1} \left[(-1)^j \frac{(n+1)!}{j!(n-j+1)!} j^{h-p-r} \right]$$

vanishes. Except for sign the quantity B_{qh} is the $(n+1)$-th difference of x^{h-p-r} for $x=0$. The degree of each difference is one less than that of the preceding difference. Therefore B_{qh} is zero when $h-p-r \leqq n$. The particular value h obtained from $h-p-r$ by putting $p=r=0$ must be included in the discussion of B_{qh}. Since N was taken as the largest value of h, it follows that $f(x)$, as given by (9), satisfies (6) if $N=n$. Moreover, the value of c_{qh} is arbitrary.

Equation (9) shows that there are $\Sigma_{h=0}^{n}(h+1) = \frac{1}{2}(n+1)(n+2)$ arbitrary constants c_{qh} which may be assigned at will. By §§ 3 and 4 it has been shown that when f is known at the $1+2+\ldots\ldots+(n+1) = \frac{1}{2}(n+1)(n+2)$ points mx_0+ky_0, m and k positive integers or zero and $m+k<n+1$, then it is known over a dense set of points covering the entire finite plane provided x_0 and y_0 are not collinear with the point zero. Since $f(x)$, as given by (9), is continuous, it is only necessary to prove that each c_{qh} is uniquely determined by assigning f at the given points mx_0+ky_0 to know that $f(x)$ is the most general continuous solution of (6) over the finite complex plane. This can be done by direct substitution,* but inasmuch as the determinant so formed is unwieldy, it is more easily accomplished by observing that the properties sought for any desired oblique network are readily deduced by a projective transformation from similar properties of a square array, on the axes of reals and imaginaries with the units 1 and i, provided only that x_0 and y_0 do not lie on the same straight line through the zero-point. Confining attention to the rectangular array mentioned, write

$$f(x) = \sum_{h=0}^{n} \sum_{q=0}^{h} A_{qh} u^{(q)} v^{(h-q)}$$

where

$$u^{(q)} = u(u-1)\ldots\ldots(u-q+1).$$

For any integral value of u less than q, $u^{(q)}=0$. Beginning with $x=0$ and proceeding outward through a triangular network similar to that employed in § 3, it is possible to determine an A_{qh} with each point of the net. Having determined the A_{qh}'s, it is only necessary to expand and collect the terms of the expression for $f(x)$ and compare with the expression involving the c_{qh}'s to completely determine each c_{qh}.

---

* The value of the determinant of the coefficients may be shown to be

$$(v_0 s_0 - u_0 t_0)^{\frac{1}{6}n(n+1)(n+2)} \prod_{a=0}^{n-1} [(n-a)!]^{2(a+1)}.$$

Thus we see that *the most general solution* $f(x) \equiv f(u+iv)$ *of equation* (6) *continuous over the finite complex x-plane is an arbitrary polynomial in u and v of degree n.*

The analytic solution of (6) over the finite complex plane is that special case of the general continuous solution for which

$$f(x) = a_0 + a_1 x + \ldots . + a_k x^k + \ldots ..$$

Replacing x by $u+iv$, it is seen at once that $a_k u^k$ must be zero for $k>n$, and hence $a_k=0$, $k>n$. Moreover, $a_k = c_{k0}$, $k \leqq n$. Hence *the most general analytic solution of* (6) *is an arbitrary polynomial in x of degree n.**

To obtain the most general continuous solution of (6) along any line in the finite complex plane, it is only necessary to observe that by the argument of §§ 3 and 4 it was proved that f is known at a dense set of points on the line if it is known at $n+1$ points of the line which are separated by some convenient unit. The argument at the beginning of this section shows that along any line not parallel to the axis of imaginaries,

$$f(x) = \sum_{j=0}^{n} a_j u^j,$$

where each a is arbitrary, satisfies (6). Since $f(x)$ is continuous, it remains only to show that the a's are determined by the functional values at the $n+1$ points on the line to know that $f(x)$ is the most general continuous solution of (6) along the line. Substitution shows immediately that the a's are uniquely determined by the $n+1$ values of f on the line. Hence *the most general solution of* (6) *continuous along any line not parallel to the axis of imaginaries is an arbitrary polynomial in u of degree n.*

Similarly, *the most general solution of* (6) *continuous along any line not parallel to the axis of reals is an arbitrary polynomial in v of degree n.*

§ 6. *Solutions of the Original Equation.*

The determination of the existence of any solution $f(x)$ of the equation

$$\sum_{i=1}^{n} \gamma_i f(\alpha_i x + \beta_i y) + \gamma_{n+1} f(x) + \gamma_{n+2} f(y) = 0, \tag{1}$$

and the determination of $f(x)$ if it exists, is accomplished by substituting the corresponding solution of the normal equation (6) in (1). It has been shown in § 5 that the general solution of (6) analytic over the finite complex x-plane

* That the analytic solution $f(x)$, if it exists, is a polynomial of degree not greater than n, is readily seen by direct differentiation of (1). If differentiation is made with respect to each of the variables y, $\alpha_j x + \beta_j y$, $j = 1, 2, \ldots, n$ it results that $f^{(k)}(y) = 0$ whenever $k > n$, because the arguments are independent in pairs which are not proportional.

or along any line in the finite complex x-plane is an arbitrary polynomial in x of degree n. Hence, the determination of the general solution of (1) analytic over the finite complex x-plane or along a line in the finite complex x-plane, if it exists, is accomplished by substituting

$$f(x)=c_0+c_1x+\ldots+c_nx^n$$

in (1). Since the result of this substitution is an identity, the coefficient of any power of the variables must vanish.

It is convenient to consider the terms of degree n independently of the remaining terms. These terms are given by

$$c_n[\sum_{i=1}^{n}\gamma_i(\alpha_ix+\beta_iy)^n+\gamma_{n+1}x^n+\gamma_{n+2}y^n]=0.$$

Now c_n is necessarily zero unless

$$[\sum_{i=1}^{n}\gamma_i(\alpha_ix+\beta_iy)^n+\gamma_{n+1}x^n+\gamma_{n+2}y^n]\equiv 0.$$

Placing the coefficients of this identity equal to zero, we have

$$\sum_{i=1}^{n}\alpha_i^n\gamma_i+\gamma_{n+1}=0,\quad \sum_{i=1}^{n}\alpha_i^{n-k}\beta_i^k\gamma_i=0,\quad \sum_{i=1}^{n}\beta_i^n\gamma_i+\gamma_{n+2}=0,\quad k=1,2,\ldots,n-1. \tag{10}$$

Since only the ratios of the γ's are significant, it implies no loss of generality to assume $\gamma_{n+2}=-1$. Under this assumption equations (10) may be employed to express the remaining γ's in terms of the α's and β's, provided $f(x)$ contains a non-vanishing term of degree n. If we write $r_i=\alpha_i/\beta_i$, then for $i<n+1$

$$\gamma_i=r_1r_2\ldots r_n/\beta_i^nr_i\prod_{h=1}^{n}{}'(r_h-r_i),$$

where the prime indicates that h does not take the value i. Solution also gives

$$\gamma_{n+1}=(-1)^nr_1r_2\ldots r_n.$$

Therefore, if c_nx^n, $c_n\neq 0$, *is a term of the analytic solution* $f(x)$ *of* (1), *it is necessary and sufficient that* (1) *may be written in the form*

$$\sum_{i=1}^{n}\frac{r_1r_2\ldots r_n}{\beta_i^nr_i\prod_{h=1}^{n}{}'(r_h-r_i)}f(\alpha_ix+\beta_iy)+(-1)^nr_1r_2\ldots r_nf(x)-f(y)=0,$$

where $r_i=\alpha_i/\beta_i$.

If the solution of (1) includes the term c_mx^m, $m<n$ and $c_m\neq 0$, then

$$\sum_{i=1}^{n}\alpha_i^m\gamma_i+\gamma_{n+1}=0,\quad \sum_{i=1}^{n}\alpha_i^{m-k}\beta_i^k\gamma_i=0,\quad \sum_{i=1}^{n}\beta_i^m\gamma_i+\gamma_{n+2}=0,\quad k=1,2,\ldots,m-1.$$

Since there are but $m+1$ linear equations in the γ's, they may be used to

express $m+1$ γ's in terms of the α's, β's and remaining γ's. For the first $m+1$ γ's in terms of the remaining quantities, solution gives

$$\gamma_i = \frac{-\sum\limits_{j=m+2}^{n} \beta_j^m \gamma_j \prod\limits_{h=1}^{m+1}{}' (r_j - r_h) - \gamma_{n+1} + (-1)^{m+1} r_1 r_2 \ldots r_{i-1} r_{i+1} \ldots r_{m+1} \gamma_{n+2}}{\beta_i^m \prod\limits_{h=1}^{m+1}{}' (r_i - r_h)}, \quad (11)$$

where the prime indicates that h does not take the value i. Hence, *if* $c_m x^m$, $m<n$ *and* $c_m \neq 0$, *is a term of the analytic solution* $f(x)$ *of* (1), *it is necessary and sufficient that each of the first* $m+1$ γ's *has the value given in* (11).

The computation involved in finding the most general solution of (1) continuous over the entire finite plane is so tedious as to make it expedient to give results only for the general second order equation. The problem for any equation is merely a matter of substitution and algebraic computation. For the second order equation

$$\gamma_1 f[(a_{11}+a_{12}i)x + (b_{11}+b_{12}i)y] \\ + \gamma_2 f[(a_{21}+a_{22}i)x + (b_{21}+b_{22}i)y] + \gamma_3 f(x) + \gamma_4 f(y) = 0,$$

where $a_{11}, a_{12}, b_{11}, b_{12}, a_{21}, a_{22}, b_{21}, b_{22}$ are real and $i=\sqrt{-1}$, the normal solution with which substitution must be made is

$$c_{00} + c_{10}u + c_{01}v + c_{20}u^2 + c_{11}uv + c_{02}v^2.$$

This substitution shows that c_{00} may be assigned different from zero when and and only when $\gamma_1+\gamma_2+\gamma_3+\gamma_4=0$. It also shows that if we write

$$A_1=\gamma_1 a_{11}+\gamma_2 a_{21}+\gamma_3, \quad B_1=\gamma_1 a_{12}+\gamma_2 a_{22}, \quad C_1=\gamma_1 b_{11}+\gamma_2 b_{21}+\gamma_4, \quad D_1=\gamma_1 b_{12}+\gamma_2 b_{22},$$

then for c_{10} and c_{01} to be independently arbitrary $A_1=B_1=C_1=D_1=0$. However, $c_{10}=c_{01}k_1$ if $A_1=B_1k_1$, $C_1=D_1k_1$, B_1 or $D_1\neq 0$, and $k_1^2=-1$. Furthermore, if we write,

$$\begin{array}{lll} A=\gamma_1 a_{11}^2+\gamma_2 a_{21}^2+\gamma_3, & B=\gamma_1 a_{11}a_{12}+\gamma_2 a_{21}a_{22}, & C=\gamma_1 a_{12}^2+\gamma_2 a_{22}^2, \\ D=\gamma_1 b_{11}^2+\gamma_2 b_{21}^2+\gamma_4, & E=\gamma_1 b_{11}b_{12}+\gamma_2 b_{21}b_{22}, & F=\gamma_1 b_{12}^2+\gamma_2 b_{22}^2, \\ G=\gamma_1 a_{11}b_{12}+\gamma_2 a_{21}b_{22}, & H=\gamma_1 a_{11}b_{11}+\gamma_2 a_{21}b_{21}, & J=\gamma_1 a_{12}b_{12}+\gamma_2 a_{22}b_{22}, \\ & K=\gamma_1 a_{12}b_{11}+\gamma_2 a_{22}b_{21}, & \end{array}$$

$c_{20}=kc_{02}$ and $c_{11}=mc_{02}$,* then it is easy to show that

(i) $k=-1$ and $m=\pm 2i$ if
  (a) $A+C$, $D+F$, $K-G$, or $H+J\neq 0$,
  (b) B, E, or $J-H\neq 0$,
  (c) $A-C=Bm$, $D-F=Em$, and $2(K+G)=(J-H)m$.

---

* The results desired here allow c_{02} to be different from zero, and no attempt is made to discuss the possibilities in case $c_{02}=0$.

(ii) $k=-1$ and $m=0$ if
  (a) A, D, H, or $G\neq 0$,
  (b) B or $E\neq 0$,
  (c) $C=A$, $F=D$, $J=H$, and $G+K=0$.

(iii) $k=-1$ and m is arbitrary if
  (a) A, D, H, or $G\neq 0$,
  (b) $B=E=0$, $C=A$, $F=D$, $J=H$, and $G+K=0$.

(iv) $k=1$ and $m=0$ if
  (a) A^2+B^2, D^2+E^2, or $G^2+H^2\neq 0$,
  (b) $C=-A$, $F=-D$, $K=G$, and $J=-H$.

The following equations, constructed on the basis of this information, answer interesting questions. The equation

$$3f[(1+2i)x+(3+i)y]-f[(1+3i)x+(6+3i)y]-5f(x)+15f(y)=0$$

has the general continuous solution $f(x)=c_{02}(u^2+v^2)$ showing that *an equation may have a continuous solution, but no non-trivial analytic solution.*

The general solution of

$$(1+i)f[3x+(1-i)y]-3f[x+2y]-(6+9i)f(x)+(10+2i)f(y)=0,$$

continuous over the finite complex x-plane is $c_{20}(u^2+2iuv-v^2)=c_{20}x^2$, showing that *an equation may have its analytic solution as the most general continuous solution.*

The elimination outlined in the footnote of § 2 shows that every solution of an equation of form (1) having no α equal to zero is included in the corresponding solution of the normal equation whose order is equal to the number of distinct ratios β_i/α_i, $i=1, 2, \ldots, n$, in the equation of form (1). One case remains, namely, that in which each ratio β_i/α_i, α_i and β_i different from zero, is equal to at least one other such ratio and at least one α and one β, of different subscripts, are zero. That there are equations of this exceptional type which have infinite series solutions is proved by the equation

$$f(x+y)-f(ix+iy)-f(-ix)-f(-iy)+f(x)+f(y)=0, \quad i=\sqrt{-1},$$

which is satisfied by a series in positive integral powers of x^4 having arbitrary coefficients.

§ 7. *The Converse Theorem.*

It was shown in § 5 that any polynomial in x of degree m_j satisfies the normal equation of order m_j. It will now be proved that any polynomial $p(x)$ in x, of degree m_j, satisfies an equation (1) whose order n is not greater than the number of non-vanishing terms of $p(x)$, plus the sum of the degrees of

such terms, and whose α's, β's and γ_{n+1} and γ_{n+2} may be assigned at will, provided a certain determinant Δ of the α's and β's is not zero as a consequence.

Suppose that

$$p(x)=a_1x^{m_1}+a_2x^{m_2}+\ldots\ldots+a_jx^{m_j}.$$

The substitution of $p(x)$ for $f(x)$ in (1) gives an identity from which

$$\sum_{i=1}^{n}\alpha_i^{m_h}\gamma_i=-\gamma_{n+1},\quad \sum_{i=1}^{n}\alpha_i^{m_h-k}\beta_i^k\gamma_i=0,\quad \sum_{i=1}^{n}\beta_i^{m_h}\gamma_i=-\gamma_{n+2}\quad k=1,2,\ldots,m-1, \tag{12}$$

for the values $h=1, 2, \ldots, j$. The total number of independent equations is not greater than $j+\Sigma_{i=1}^{j}m_i$. Setting n equal to the number of independent equations, we have a system of non-homogeneous linear equations in n unknowns $\gamma_1, \gamma_2, \ldots, \gamma_n$, provided γ_{n+1} and γ_{n+2} are not both assigned equal to zero. A necessary and sufficient condition that the unknown γ's are uniquely determined in terms of the α's, β's and two assigned γ's is that the determinant Δ of the coefficients be different from zero. It is obvious that Δ is a polynomial in the α's and β's. Furthermore, it is at once evident that the term formed by the product of the elements in the principal diagonal is unique. Hence the polynomial in the α's and β's does not vanish identically. Therefore the α's and β's may be assigned in any way such that the polynomial Δ has a value different from zero.

If γ_{n+1} and γ_{n+2} are both assigned equal to zero, equations (12) form a system of n linear homogeneous equations in n unknowns, and the non-vanishing of Δ is a necessary and sufficient condition that each of the unknown γ's is zero. In this case the equation is trivially satisfied by $p(x)$. In general, then, *any polynomial $p(x)$ satisfies an infinity of equations* (1) *whose orders do not exceed the number of non-vanishing terms of $p(x)$ plus the sum of the degrees of such terms and whose α's, β's, γ_{n+1}'s and γ_{n+2}'s may be assigned at will, provided only that a polynomial Δ of the α's and β's does not vanish for the assigned values.*

§ 8. *Discontinuous Solutions.*

Any solution $f(x)$ of an equation (1) satisfies the normal equation whose order n is determined by (1). Suppose that the domain of $f(x)$ is any line in the finite complex x-plane, and that $f(x)$ is continuous in some interval of length $\delta>0$ of the line. It is readily seen from the normal equation satisfied by f that if the first $n+1$ arguments are so chosen that they represent points in the interval while the remaining argument represents a point outside the interval, $f(x)$ is determined at the last-named point as the sum of continuous functions. Therefore f is continuous at the outside point. In this way f may be shown to be continuous at all points in the two intervals of length

δ/n which lie at the ends of the given interval. Therefore f is continuous in an interval of length $[(n+2)\delta]/n$. Any finite interval of length σ may be reached in this manner by a finite number of extensions of the interval of length δ. We may therefore state that *if $f(x)$ has a finite point of discontinuity on any line in the complex x-plane, it has a point of discontinuity in every interval of the line, however small.*

Suppose $f(x)$ is continuous in a region of the finite complex x-plane. A circle may be inscribed in the region such that $f(x)$ is continuous in the closed region of which the circle is the boundary. Suppose the radius of this circle is δ and consider a concentric circle of radius $[(n+1)\delta]/n$. By means of the normal equation f may be determined at any point in the area between the circles as the sum of $n+1$ continuous functions, namely, f at $n+1$ points in the circle of radius δ. Therefore f at the point between the circles is continuous. This is true for every point of the area between the circles, and hence f is continuous in the circle of radius $[(n+1)\delta]/n$. This process may obviously be repeated to prove that f is continuous in any finite region of the plane. Hence, *if $f(x)$ has a point of discontinuity in the finite complex x-plane, it has a point of discontinuity in every finite region of the plane.*

G. Hamel (*loc. cit.*) has exhibited a discontinuous solution* $f(x)$ of the Cauchy equation

$$f(x+y)=f(x)+f(y).$$

From the treatment in § 2 it is clear that f also satisfies

$$f(2x+y)-2f(x+y)+f(y)=0.$$

Replacing y by $hx+y$, multiplying the equation by $(-1)^h(n-1)!/h!(n-h-1)!$ for successive values $h=0,1,\ldots,n-1$, and adding the n equations so formed, it is easily seen that

$$\sum_{k=0}^{n+1}(-1)^k\frac{(n+1)!}{k!(n+1-k)!}f(kx+y)=0. \tag{6}$$

We therefore have a discontinuous solution of the normal equation for each order n.

§ 9. *Certain Types of Equations having Variable Coefficients.*

The functional equations that have been discussed may be employed to solve certain equations of the form,

$$\sum_{i=1}^{n}\phi_i(x,y)f(\alpha_i x+\beta_i y)+\phi_{n+1}(x,y)f(x)+\phi_{n+2}(x,y)f(y)+\phi_{n+3}(x,y)=0, \tag{13}$$

where the ϕ's are known functions. A general statement and a few examples suffice to indicate some of the equations that may be solved. Suppose there are k transformations

$$x=\gamma_j x'+\delta_j y', \quad y=\lambda_j x'+\mu_j y',$$

* This solution is obtained on the assumption of the validity of the Zermelo axiom.

which may be applied to (13) to obtain new equations such that if each equation is multiplied by a non-zero constant, the sum of them is of form (1). A solution of (13), if it exists, is included in the corresponding solution of the auxiliary equation of form (1) deduced from (13). In order to find a solution of (13) it is sufficient to substitute the solution of the auxiliary equation and compute the coefficients of the variables.

Equation (13) includes the non-homogeneous equation in which $\phi_i(x, y)$, $i=1, 2, \ldots, n+2$, is further restricted to be a constant. In this case it is obvious from equation (13) that $\phi_{n+3}(x, y)$ is a polynomial in x and y if an analytic solution exists, and a polynomial in u, v, s and t if a continuous solution exists. Such a non-homogeneous equation is*

$$f(x+y)=f(x)+f(y)+2xy. \tag{14}$$

Transformations which may be used to solve this equation are

$$x=x',\ y=x'-y' \quad \text{and} \quad x=x',\ y=y'-x'$$

whence, after dropping the primes,

$$f(2x-y)-f(x-y)-f(-x+y)-2f(x)+f(y)=0.$$

The arguments $x-y$ and $-x+y$ are proportional and the normal equation is therefore of order 2. Hence the general solution of (14), analytic over the finite complex x-plane is readily seen to be $f(x)=a_1x+x^2$, where a_1 is arbitrary. The general solution of (14) continuous over the finite complex x-plane is $f(x)=a_{10}u+a_{01}v+x^2$, where a_{10} and a_{01} are arbitrary.

Suppose all the ϕ's are constant and $\phi_{n+3}\neq 0$. If $\Sigma_{i=1}^{n+2}\phi_i\neq 0$, then $f(0)$ is finite and uniquely determined, and the transformation $f(x)=g(x)+f(0)$ may be employed to obtain an equation (1) of order n in $g(x)$. Hence $f(x)$ is a polynomial (in x, in u and v, in u or in v, as the case may be) of degree not greater than n. The transformation used in this case has the advantage of furnishing an auxiliary equation (1) whose order is not greater than that of the original equation (13).

An equation which illustrates reduction by interchanging arguments of f is

$$(\cos^2 x)f(x+y)+(\sin^2 x)f(x-y)-f(x)-f(y)-2(\cos 2x)xy=0.$$

If y is replaced by $-y$ the equation becomes

$$(\sin^2 x)f(x+y)+(\cos^2 x)f(x-y)-f(x)-f(-y)+2(\cos 2x)xy=0.$$

The sum of these equations is of form (1). The solution analytic over the finite complex x-plane of the equation having variable coefficients is $f(x)=x^2$. It is easily seen that this is the most general continuous solution.

The equation

$$2f(2x+y)+(x+y)f(x-y)+3f(x)-3f(y)=0$$

---

* *American Mathematical Monthly*, Vol. XXIV (1917), p. 178.

may be reduced by replacing x by $2x+y$ and y by $x+2y$, whence

$$2f(5x+4y)+3(x+y)f(x-y)+3f(2x+y)-3f(x+2y)=0,$$

and subtracting 3 times the original equation from the transformed equation. The reduced equation is

$$2f(5x+4y)-3f(2x+y)-3f(x+2y)-9f(x)+9f(y)=0.$$

It is readily seen that the equation with variable coefficient has no continuous solution.

If one of the first $n+2$ ϕ's of (13) is variable while the remaining ϕ's are constant the product of that ϕ and the corresponding $f(\alpha_i x+\beta_i y)$ is a polynomial. The variable ϕ is therefore a rational function whose denominator is a factor of $f(\alpha_i x+\beta_i y)$. An equation illustrating this point is

$$2\left[\frac{x+y}{x-y}\right]^2 f(x-y)-f(x+2y)-f(x)+2f(y)=0.$$

The equation obtained by replacing x by $2x+y$ and y by $x+2y$ is

$$18\left[\frac{x+y}{x-y}\right]^2 f(x-y)-f(4x+5y)-f(2x+y)+2f(x+2y)=0.$$

The equation obtained by subtracting the transformed equation from 9 times the original one gives an equation (1) of order 3. The analytic solution $f(x)$ of the equation with a variable coefficient is then easily seen to be ax^2, where a is arbitrary.

§ 10. *Application to Binomial Equations.*

An application of linear functional equations having constant coefficients may also be made to certain equations of the form,

$$\prod_{i=1}^{k}[f(\alpha_i x+\beta_i y)]^{\gamma_i}=C\prod_{i=k+1}^{n+1}[f(\alpha_i x+\beta_i y)]^{\gamma_i}[f(y)]^{\gamma_{n+2}}, \tag{15}$$

where C is a constant, the real part of each γ is positive and no α is zero. Let us first consider the solution $f(x)$ of (15) analytic at all points in the finite complex plane. Suppose that $f(x)$ has a zero at some point $x=a$. Let $y=a$ in (15). Then

$$\prod_{i=1}^{k}[f(\alpha_i x+\beta_i a)]^{\gamma_i}=0$$

for all values of x. Hence there is a finite region in which $f(x)$ has an infinity of zeros. But this is impossible since $f(x)$ is analytic throughout the finite plane. Therefore, *when $f(x)$ is analytic throughout the finite complex plane, and not identically zero, it is never zero.* The case of a continuous solution $f(x)$ of (15) presents more difficulty. It is evident that *if one member of* (15) *has either no factor or only one factor involving f, then $f(x)$ is never zero unless it is identically so.*

The function $\phi(x)=\log f(x)$, where it is understood that the principal determination of the logarithm is employed, is analytic or continuous with f when the latter has no zeros. Therefore, in each of the cases considered above, $\phi(x)$ satisfies the equation

$$\sum_{i=1}^{k}\gamma_i\phi(\alpha_i x+\beta_i y)-\sum_{i=k+1}^{n+1}\gamma_i\phi(\alpha_i x+\beta_i y)-\gamma_{n+2}\phi(y)-K+2s\pi i=0,$$

where K is the principal determination of $\log C$ and s is any integer. For any given s $\phi(x)$ is a polynomial of degree not greater than n. Furthermore, the argumentation of § 9 shows that a variation in s affects only the constant term of $\phi(x)$. Therefore, *in each of the cases considered above, $f(x)$ is an exponential function of the form*

$$f(x)=e^{P+k(s)}, \tag{16}$$

where P is a polynomial of degree not greater than n, and $k(s)$ is a constant depending on s. Thus we see that in general the solutions of (15) are given by (16) for the various possible values of $k(s)$.

The same result may be stated for the continuous solution $f(x)$ of (15), when each member involves at least two factors containing f, provided $f(0)\neq 0$. For $\phi(x)$ as defined, is continuous in some region about the point $x=0$ since $f(x)$ is different from zero in some such region. For a given s, therefore, $\phi(x)$ must be everywhere continuous in the finite complex plane because it satisfies a non-exceptional equation of type (1).

The equation,

$$\psi(x+y)\psi(x-y)=[\psi(x)\psi(y)]^2,$$

mentioned in § 1, is included in the last case considered. For suppose there is at least one point $x=b$ at which $\psi(x)$ is not zero. Let $x=y=b$. Then

$$\psi(2b)\psi(0)=[\psi(b)]^4\neq 0,$$

and $\psi(0)\neq 0$. It is easily seen that the solutions $\psi(x)$ analytic over the finite complex plane are

$$\psi(x)=e^{ax^2-s\pi i}=\pm e^{ax^2},$$

and the solutions $\psi(x)$ continuous over the finite complex plane are

$$\psi(x)=e^{c_{20}u^2+c_{11}uv+c_{02}v^2-s\pi i}=\pm e^{c_{20}u^2+c_{11}uv+c_{02}v^2},$$

where a and the c's are arbitrary constants.

§ 11. *Equations Involving More than One Function.*

Consider the equation

$$\sum_{i=1}^{n}\gamma_i f_i(\alpha_i x+\beta_i y)+\gamma_{n+1}f_{n+1}(x)+\gamma_{n+2}f_{n+2}(y)=0, \tag{17}$$

where no γ is zero and the f's are unknown, continuous, single-valued functions to be determined if possible so that (17) shall be identically satisfied by them. The functions f_i may or may not all be distinct. The method of elimination employed in § 2 is applicable to (17). If no α is zero it is evident then that f_{n+2} satisfies the normal equation of order n. Therefore f_{n+2} is a polynomial of degree not greater than n. Under the assumption that no α_i and no β_i are zero, and no two of the ratios β_i/α_i are equal, any term may be given the argument y by a linear transformation which makes no α and no β zero. In this case, therefore, *every function f_i of (17) is a polynomial of degree not greater than n.* To find any necessary restrictions on the coefficients of these polynomials, it is sufficient to substitute the n-th degree polynomials having general coefficients in (17), and to equate to zero the resulting coefficients of the variables. It is obvious that equation (1) is a special case of (17).

If some of the ratios β_i/α_i are equal it may be assumed without loss of generality that equation (17) is so arranged that functions having arguments of a common ratio are placed consecutively. If all the functions having arguments of a common ratio have subscripts i, such that $g \overline{\gtrless} i \overline{\gtrless} h$, then we may write

$$F_h(\alpha_h x+\beta_h y)=\sum_{i=g}^{h}\gamma_i f_i(\alpha_i x+\beta_i y).$$

Equation (17) may now be written

$$\Sigma F_h(\alpha_h x+\beta_h y)+F_{n+1}(x)+F_{n+2}(y)=0, \tag{18}$$

where no α and no β are zero, and no two ratios β_h/α_h are equal. We denote by $q+2$ the number of terms in the first member of (18). Each F is a polynomial of degree not greater than q. Each F therefore determines a non-homogeneous equation in certain f's having arguments differing by constant factors. If the f's of any F are identical, that is, if the f's of F are the same function, F determines a non-homogeneous mixed q-difference equation satisfied by f. The equations of type (1), which have no α equal to zero, but have some ratios β_i/α_i equal, or some β's zero, are special cases of (18) given by $F_{n+2}(y)=\gamma_{n+2}f(y)$. The equations of the exceptional case noted in the last paragraph of § 6 are equations of form (18) which have no F a constant multiple of a single f. In this connection it is interesting to note that the function f of the example in the paragraph cited satisfies the equations

$$f(x)-f(-ix)=f(x)-f(ix)=0,$$

whence

$$f(-ix)=f(ix)=f(-x)=f(x).$$

*Contributions to the Study of Oscillation Properties of the Solutions of Linear Differential Equations of the Second Order.**

BY R. G. D. RICHARDSON.

Introduction.

The study of boundary problems for linear differential equations of the second order dates back to the time of Euler and D'Alembert, with whom it arose in connection with problems of mathematical physics. Beginning with the fundamental paper of Sturm in 1836, there have been extensive investigations † in this field in recent years, notably by Klein, Bôcher, Stekeloff, Kneser, Hilbert and Birkhoff. Since the differential equation of the second order is of such fundamental importance in so many fields, and since similar general problems for equations of higher order can not be handled by processes so far devised, the invention of new methods and further investigation of the nature of solutions find ready justification.

The equation to be studied will be taken in the form

$$\frac{d}{dx}\left(p(x)\frac{dy(x)}{dx}\right)+G(x,\lambda)y(x)\equiv(py_x)_x+G(x,\lambda)y=0,\qquad 0\leqq x\leqq 1,\qquad (1)$$

where $G(x, \lambda)$ is a function depending on a parameter λ. The solution $y(x)$ of this self-adjoint equation shall be subject to the self-adjoint boundary conditions

$$\left.\begin{aligned}&\alpha_1 y(0)+\alpha_2 y_x(0)+\alpha_3 y(1)+\alpha_4 y_x(1)=0,\\&\beta_1 y(0)+\beta_2 y_x(0)+\beta_3 y(1)+\beta_4 y_x(1)=0,\\&p(1)(\alpha_1\beta_2-\alpha_2\beta_1)=p(0)(\alpha_3\beta_4-\alpha_4\beta_3),\end{aligned}\right\}\qquad (2)$$

where the two sets of real coefficients $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ and $\beta_1, \beta_2, \beta_3, \beta_4$ are linearly independent. The most important special cases of these boundary conditions are given by

$$y(0)=y(1)=0;\quad y(0)=y_x(1)=0;\quad y_x(0)=y(1)=0;\quad y_x(0)=y_x(1)=0.\qquad (2')$$

---

* Read before the American Mathematical Society, September 4, 1917.

† For existing methods and literature of the subject see Bôcher, *Encyklopädie Mathematischen Wissenschaften*, II A7a, *Proceedings International Congress of Mathematicians*, Vol. I (1912), p. 163, and "Leçons sur les Methodes de Sturm" (1917); Lichtenstein, *Rendiconti del Circolo Matematico di Palermo*, Vol. XXXVIII, p. 113.

There are many interesting questions in regard to this linear problem. Do there exist parameter values λ such that there are solutions of (1) satisfying relations (2)? If so, how many are there, and how are they distributed? What is the nature of the corresponding solutions? How do the solutions vary with change of the coefficients of the equation and of the boundary conditions? When does the totality of solutions form a fundamental set in terms of which functions may be expanded?

Among the methods which have been used in studying the problem are: (1) Differential equations including the use of comparison, approximation and asymptotic expressions; (2) the minimum principle in the calculus of variations; (3) integral equations; (4) the theory of linear algebraic equations in an infinite number of variables;* (5) a limiting process with linear algebraic approximating difference equations (cf. § 1, IV). The methods developed in this memoir would fall under (1) and (2) and may be characterized as a free use of the differentiation of fundamental formulae with regard to the parameters involved in the equations and in the boundary conditions, together with the proof that under the conditions imposed certain integrals are positive.

Exact oscillation theorems for solutions under the boundary conditions (2) have been developed by Birkhoff † for the case of the special equation

$$y_{xx}+G(x,\lambda)y=0, \quad \frac{\partial G}{\partial \lambda}>0, \quad \lim_{\lambda=-\infty} G=-\infty, \quad \lim_{\lambda=+\infty} G=+\infty.$$

Another important special case of equation (1) in which $G(x,\lambda)$ contains the parameter linearly

$$(py_x)_x+(q+\lambda k)y=0, \tag{3}$$

has been studied very extensively. When $k(x)>0$ this equation may be reduced to a form included in that investigated by Birkhoff.

The *definite* case of (3), viz., when one of the integrals

$$\int_0^1 ky^2dx, \quad \int_0^1 (py_x^2-qy^2)\,dx$$

has one sign for all functions $y(x)$ considered, has been discussed in many phases by mathematicians since the time of Sturm. By means of his theory of integral equations Hilbert ‡ established the *existence* of characteristic

---

* Lichtenstein, *loc. cit.*

† *Transactions of the American Mathematical Society*, Vol. X (1909), p. 259. It should be remarked that both in the second and third lines from the end of the statement of the principal theorem (p. 260) instead of $p+1$ we should read $p-1$.

‡ "Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen" (Teubner, 1912).

parameter values and characteristic solutions of equation (3) for this definite case. As will be shown in § 5 the boundary conditions which he used are normal forms of (2). Oscillation theorems for equation (3), under some of the simple boundary conditions (2′), have been established by various means; among others by setting up the corresponding calculus of variations problem, and interpreting the Jacobi criterion.*

When k has both signs and q is positive in at least a part of the interval and sufficiently large there, it is not necessary that either of the integrals in question be definite. This case, which we shall call the *non-definite*, was first discussed incidentally by the author in a paper † in which he was treating the problem of oscillation theorems for two equations with two parameters. He showed that when the boundary conditions are $y(0)=y(1)=0$, there exists an integer n_1 such that for $n<n_1$ there are *no real solutions* which have n zeros, while for $n>n_1$ there are *at least two.*‡ At that time all the principal results of §§ 2–4 were obtained, but were not published.

For the special equation (3) and $k>0$ the theorems proved by Birkhoff were rediscovered by Haupt § in his dissertation. The oscillation theorem stated in this dissertation for the case that k changes sign is corrected in a later article,‖ and by using the methods which I had developed, various oscillation theorems for the general equation (1) are derived. It is also shown by means of expansion theorems that if the non-definite equation (3) be taken in a certain normal form there exists an integer n_2 such that for $n>n_2$ there are precisely two solutions with n zeros and satisfying the boundary conditions (2).

The object of the present memoir is to investigate the conditions to be imposed on $G(x, \lambda)$ regarded as a function of λ, so that definite oscillation theorems for solution of (1) may be determined. In § 2 an attempt is made to bring as close together as possible necessary conditions and sufficient conditions for a limited or unlimited number of oscillations. A criterion for the behavior of the zeros with change of the parameter is obtained in § 3. This permits the development of very general theorems for the unique existence of solutions vanishing at the end points and possessing a prescribed number of zeros, and also for the existence of two, and only two solutions of this nature.

---

* *Mathematische Annalen*, Vol. LXVIII (1910), p. 279.

† *Transactions of the American Mathematical Society*, Vol. XIII (1912), p. 22.

‡ That there were *exactly* two when $n>n_1$ was stated by the author in a paper in the *Mathematische Annalen*, Vol. LXXIII, p. 289, in which he was discussing oscillation theorems for three linear equations with three parameters. The subsequent results of the memoir were not affected by this error which was corrected by a note in Vol. LXXIV (1913), p. 312, of the same journal.

§ "Untersuchungen über Oszillationstheoreme" (Teubner, 1911).

‖ Haupt, *Mathematische Annalen*, Vol. LXXVI, p. 67.

These theorems contain as special cases all known results in this field, and some new special cases are set forth in detail.

The non-definite case of (3) and (1) is discussed in § 4. The question of whether there may be for a given oscillation number more than two parameter values for which there are solutions of (3) (2) is settled by giving an example in which for any n in an interval n_1, n_2 there are four values of λ corresponding. That there exists an integer n_2 such that for $n>n_2$, there are exactly two solutions is proved by a method entirely different from that of Haupt, and in some particulars it would seem that the resulting theorem is less satisfactory, in others more satisfactory than his. The theory is also extended to cover some corresponding cases of (1). Concerning the complex solutions which correspond to values of $n<n_1$ some theorems are derived.

In the latter half of the memoir a method is developed for obtaining the facts in regard to the solutions of the *general* equation (1) under the general boundary conditions (2). With this end in view § 5 is devoted to a reduction of the boundary conditions to normal forms by means of the usual transformation of the dependent variable y, which leaves the number of zeros unchanged. Each of these three normal forms,

I. $\sigma y(0)+y_x(0)=0, \qquad \tau y(1)+y_x(1)=0;$

II. $y(0)=hy(1), \qquad hp(0)y_x(0)=p(1)y_x(1);$

III. $y(0)=lp(1)y_x(0), \qquad lp(0)y_x(0)=y(1)$

contains one or two parameters σ, τ, h, l. When special values 0 or ∞ are assigned to these parameters, the forms reduce to the simple cases (2′) for which the facts are readily obtainable from the developments of the earlier sections. The first form is of essentially different character from the others. In the latter the parameter h or l is a double-valued function of λ, which in general is real in sub-intervals only, while in the former each of the parameters σ, τ is a single-valued function of the other and of λ, and is real throughout. By letting λ vary, and calculating the rates of change of these parameters and of $G(x, \lambda)$, theorems of oscillation are derived (§§ 6–8) for boundary conditions in each of the normal forms. While detailed results are not given in all cases, these are immediate developments of the fundamental facts ascertained.

Any linear differential equation of the second order

$$\phi y_{xx}+\psi y_x+\theta y=0 \quad \text{or} \quad \phi y_{xx}+\psi y_x+(\lambda\theta_1+\theta_2)y=0, \quad \phi>0,$$

may be thrown into the corresponding self-adjoint form (1) or (3) on multiplying by the function $p(x)=e^{\int_0^x \frac{\psi}{\phi}dx}$; the corresponding self-adjoint boundary

conditions (2) are changed in form only by the substitution $p(0)=1$, $p(1)=e^{\int_0^1 \frac{\psi}{\phi} dx}$.

The oscillation theorems remain unchanged if the variables are subjected to the usual transformations of the dependent and independent variables $y=\eta\bar{y}$, $x=\xi(\bar{x})$ $\left(\eta(x)\neq 0, \frac{d\xi}{d\bar{x}}\neq 0\right)$. The resulting equation is of the form

$$\bar{\phi}\bar{y}_{\bar{x}\bar{x}}+\bar{\psi}\bar{y}_{\bar{x}}+\bar{\theta}\bar{y}=0 \quad \text{or} \quad \bar{\phi}\bar{y}_{\bar{x}\bar{x}}+\bar{\psi}\bar{y}_{\bar{x}}+(\lambda\bar{\theta}_1+\bar{\theta}_2)\bar{y}=0,$$

the new boundary conditions for the new interval $\bar{x}_0$, $\bar{x}_1$ being of the form

$$\begin{aligned}
&\bar{\alpha}_1\bar{y}(\bar{x}_0)+\bar{\alpha}_2\bar{y}_{\bar{x}}(\bar{x}_0)+\bar{\alpha}_3\bar{y}(\bar{x}_1)+\bar{\alpha}_4\bar{y}_{\bar{x}}(\bar{x}_1)=0,\\
&\bar{\beta}_1\bar{y}(\bar{x}_0)+\bar{\beta}_2\bar{y}_{\bar{x}}(\bar{x}_0)+\bar{\beta}_3\bar{y}(\bar{x}_1)+\bar{\beta}_4\bar{y}_{\bar{x}}(\bar{x}_1)=0,\\
&e^{\int_{\bar{x}_1}^{\bar{x}_2}\frac{\bar{\psi}}{\bar{\phi}}d\bar{x}}[\bar{\alpha}_1\bar{\beta}_2-\bar{\alpha}_2\bar{\beta}_1]=\bar{\alpha}_3\bar{\beta}_4-\bar{\alpha}_4\bar{\beta}_3,
\end{aligned}$$

as may be shown by computation.* The invariantive property of self-adjointness gives to the results obtained for (1) or (3) a very general character.

§ 1. *Some Properties of Solutions of the Differential Equation.*

In his original memoir Sturm studied the differential equation in the form

$$\frac{d}{dx}\left(K(x,\lambda)\frac{dy}{dx}\right)+\bar{G}(x,\lambda)y=0,$$

where $K>0$ and $\bar{G}$ both depend on a parameter λ. But by a change of variables this may be reduced to the form

$$(py_x)_x+G(x,\lambda)y=0, \tag{4}$$

where $p(x)$ is a positive function independent of λ. We lose nothing in generality by considering this latter equation. Regarded as functions of x the coefficients p, G will be postulated as continuous together with as many derivatives as is desired, while G will be considered as analytic with respect to λ. The usual modifications of the results derived can be written down immediately if less stringent hypotheses are imposed. The trivial solution $y\equiv 0$ will be excluded from the discussion. Some theorems concerning solutions of (4) will now be reviewed.

I. If one zero of a solution of (4) is held fixed, all others are moved nearer to it by a decrease of p or an increase of G. If, for example, $p(x)$ is less than a constant P, and $G(x,\lambda)$ is greater than a constant $\gamma>0$, the zeros of (4) are closer together than those of the equation

$$y_{xx}+gy=0, \quad g=\frac{\gamma}{P},$$

* For the special case of a transformation of the dependent variable only, this computation is given in § 5.

which has a solution $y=\sin\sqrt{g}(x+c)$ with zeros at intervals of $\frac{\pi}{\sqrt{g}}$. By fixing p, and taking G large enough in any interval of x, the zeros of (4) may then be made as close together as desired.

II. The special equation where $G(x, \lambda)$ contains the parameter linearly,

$$(py_x)_x+(q+\lambda k)y=0, \tag{5}$$

has been much discussed. The boundary conditions

$$y(0)=y(1)=0 \tag{6}$$

are of the greatest interest and three cases may be distinguished.

(A) *Orthogonal Case*, when $k(x)\geqq 0$. There is then an infinite number of parameter values $\lambda_m(\lambda_1\leqq\lambda_2\leqq\ldots.)$ with a limiting point at positive infinity only, for each of which a solution Y_m satisfying (6) exists. The number of zeros of the solution Y_m (including those at $x=0$ and $x=1$) is $m+1$.

(B) *Polar Case*, when $k(x)$ takes on both signs and the integral

$$D(y)=\int_0^1(py_x^2-qy^2)\,dx \tag{7}$$

is positive-definite,* that is, for the given boundary condition (6) $D(y)$ can not take on negative values. There are two sets each of an infinite number of parameter values $0\leqq\lambda_1\leqq\lambda_2\ldots.$, $0\geqq\lambda_{-1}\geqq\lambda_{-2}\ldots.$, with limiting points at positive and negative infinity respectively, corresponding to which solutions Y_m, Y_{-m} exist. For both Y_m and Y_{-m} the number of zeros is $m+1$.

(C) *Non-definite Case*, when both of the integrals

$$\int_0^1 ky^2dx, \quad \int_0^1(py_x^2-qy^2)\,dx$$

may take on negative values. This will be discussed in § 4.

III. For the equation (5) and the special boundary conditions (6) certain minimum properties may be stated. In the orthogonal case the minimum of the integral $D(y)$ for those values of y which satisfy (6), and the normalizing and orthogonalizing conditions

$$\int_0^1 ky^2dx=1, \quad \int_0^1 kY_iy\,dx=0, \qquad i=1, 2, \ldots., m-1, \tag{8}$$

is λ_m, and is furnished by the normalized solution Y_m of (5). In the polar case the minimum of $D(y)$ for those values which satisfy conditions (6) and (8) is λ_m and is furnished by Y_m; the minimum subject to the conditions (6) and

$$\int_0^1 ky^2dx=-1, \quad \int_0^1 kY_{-i}y\,dx=0, \qquad i=1, 2, \ldots., m-1,$$

* Bôcher has pointed out (*Proceedings International Congress of Mathematics, loc. cit.*, p. 173) that the special case of the polar problem where $q\leqq 0$ can be treated by the method of Sturm. This remark, however, does not apply to the polar case in its most general form.

is $-\lambda_{-m}$ and is furnished by Y_{-m}. In the non-definite case, $D(y)$ can be negative and the minimum (even in the simplest problem ($m=1$)) may not exist. However, in view of the developments of § 4, it would seem probable that by a modification of the discussion, the solution Y_m, for m large enough, may be regarded as furnishing a minimum of a calculus of variations problem.

IV. While the notion of regarding a differential equation directly as the limit of a set of difference equations has been used heuristically since the time of Euler, it was, so far as the author is aware, first made definite and rigorous in a recent paper.* In that paper the problem actually discussed is that of existence theorems for partial differential equations with given boundary conditions. But the same method applies, for example, to equation (1), and the discussion is essentially simpler for the equation in one dimension than for that in two or more. For the sake of simplicity let us confine ourselves to the case (5) (6) and consider the unit interval to be divided into m equal parts, the values of y, p, q, k at the point $\frac{i}{m}(i=0, 1, \ldots, m)$ to be denoted by y_i, p_i, q_i, k_i, and difference equations

$$m^2[p_{i+1}(y_{i+1}-y_i)-p_i(y_i-y_{i-1})]+q_iy_i+\lambda k_iy_i=0 \quad (i=1, 2, \ldots, m-1), \tag{9}$$

to be set up. In order that there be solutions of these equations λ must be one of the $m-1$ roots of the determinant formed from the coefficients. With increase of m the number of points at which y is defined increases, but we can pick out corresponding parameter values and solutions of the various sets of difference equations, and if proper continuity conditions are imposed on the coefficients of the differential equations, it may be shown that the corresponding sets of parameter values approach as a limit a parameter value of (5), and corresponding solutions approach a solution of (5). In this way the infinite set of solutions of the differential equation is obtained. If k has both signs and q is positive and sufficiently large, at least in some part of the interval, some of the λ's and the corresponding y's will be complex.† But in all those cases of equation (5) heretofore treated (the orthogonal and polar), the method of passing to the limit in (9) suffices and gives a simple proof of the fundamental facts. This method can be extended to a treatment of existence theorems for solutions of the equation (4) with more general boundary conditions.

---

* *Transactions of the American Mathematical Society*, Vol. XVIII (1917), p. 489.

† One method of proving this would be by noting that the parameter values and solutions are approximations to those of the differential equation which are shown in § 4 to be complex.

§ 2. *Sufficient Conditions for the Existence of Solutions with an Unlimited Number of Zeros.*

In the consideration of the equation

$$(py_x)_x+G(x,\lambda)y=0 \tag{10}$$

let us impose the restriction that for finite values of λ the function $G(x,\lambda)$ is limited, and in particular that $G(x,0)$ is limited. The interesting cases will be covered by one of two hypotheses, which will be justified by the later developments of this section.

HYPOTHESIS A. *For at least one point of the interval, the upper limit of $G(x,\lambda)$ becomes infinite with λ ($\overline{\lim\limits_{\lambda=+\infty}} G(x,\lambda)=+\infty$), and in such a manner that the number of zeros of the solutions increases without limit with λ.*

The problem treated by Birkhoff* where $\frac{\partial G}{\partial \lambda}>0$ and $G(x,+\infty)=+\infty$ is a special case, and the orthogonal problem (§1, II) is still more special.

HYPOTHESIS B. *For at least one point of the interval $\overline{\lim\limits_{\lambda=+\infty}} G(x,\lambda)=+\infty$; for at least one other, $\overline{\lim\limits_{\lambda=-\infty}} G(x,\lambda)=+\infty$; and in both cases $G(x,\lambda)$ increases in such a manner that the number of zeros of the solutions increases without limit with λ.*

The polar case (§1, II) is included in this hypothesis.

THEOREM I. *In order that there be a set of parameter values λ such that the number of oscillations of the corresponding solutions of (10) be unlimited, it is necessary that in the neighborhood of at least one point $\overline{\lim\limits_{\lambda=+\infty}} G(x,\lambda)=+\infty$ or $\overline{\lim\limits_{\lambda=-\infty}} G(x,\lambda)=+\infty$.*

For, as the number of zeros increases, the length of the smallest interval decreases without limit. The zeros of the equation

$$Py_{xx}+G(x,\lambda)y=0, \quad P=\text{maximum } p(x) \tag{11}$$

are farther apart than those of (10) (§ 1, I). To establish the theorem for (10) it is then only necessary to prove it for (11). Let us denote by α, β that pair of consecutive zeros of (11) whose distance is a minimum. In such an interval α, β $[y(\alpha)=y(\beta)=0]$, y may be taken positive, and since the equation is homogeneous, its solution for all values of λ may be multiplied by a constant so that the maximum is 1. The maximum of y_x must be at least as great as $\frac{1}{\beta-\alpha}$ which is the slope of the line joining $(\alpha,0)$, $(\beta,1)$. To investi-

* *Loc. cit.*

gate the maximum of $-\frac{y_{xx}}{y}$ we note that y_x is zero at some point of the interval and at least as great as $\frac{1}{\beta-\alpha}$ at another. Hence $-y_{xx}$ must be at least as great as $\frac{1}{(\beta-\alpha)^2}$ at some point, and further $\max\left(-\frac{y_{xx}}{y}\right)\geqq\frac{1}{(\beta-\alpha)^2}$. Since for at least one sub-interval the length approaches zero, and since

$$G(x,\lambda)=-\frac{Py_{xx}}{y},$$

it follows that for the interval 0, 1 $\underline{\lim}\max G(x,\lambda)=+\infty$. Since by hypothesis G can become infinite only for $\lambda=\pm\infty$ the theorem may be readily deduced from these results by the usual processes of reasoning.

To show that the necessary condition of *Theorem* I is not sufficient, let us consider the following example: In the interval 0, $\frac{1}{2}-\epsilon$ we set up the function $y=\frac{2}{\pi}\sin\frac{\pi x}{2}$, and in the interval $\frac{1}{2}+\epsilon$, 1 the function $y=\frac{2}{\pi}\sin\frac{\pi}{2}(1-x)$; both of these arcs satisfy the equation $y_{xx}+\frac{\pi^2}{4}y=0$. If continued, they would meet at the point $x=\frac{1}{2}$ with an angle arc tan $2\sqrt{2}$. If in the interval $\frac{1}{2}-\epsilon$, $\frac{1}{2}+\epsilon$ an analytic curve is introduced which is tangent to these two arcs $\left(\text{on one of which } y_x>\frac{1}{\sqrt{2}} \text{ and on the other } y_x<-\frac{1}{\sqrt{2}}\right)$, it must have at some point $-y_{xx}>\frac{1}{\sqrt{2}\epsilon}$, and if ϵ is taken small enough, $\max\left(-\frac{y_{xx}}{y}\right)$ is great at pleasure. Hence denoting by $y_{xx}+G(x,\lambda)y=0$ the differential equation which has for solution the function defined for the interval 0, 1 by this method, and setting $\lambda=\frac{1}{\epsilon}$, we know that $\lim\limits_{\epsilon=0}\max G(x,\lambda)$ in the neighborhood of $x=\frac{1}{2}$ increases without limit. On the other hand none of the suite of functions y has any zeros in the interval.

Having shown by this example that the necessary condition of *Theorem* I is not sufficient, let us now deduce a sufficient criterion that when $\lim\limits_{\lambda=\infty}G(x,\lambda)=\infty$, the number of zeros of solutions of (10) be unlimited. A similar discussion can be given for the case $\lim\limits_{\lambda=-\infty}G(x,\lambda)=\infty$. Denoting as before by P the maximum of $p(x)$, and by M an arbitrarily large constant, it is possible to find a λ and an x-interval such that $G(x,\lambda)>MP$ in this interval, whose length will be denoted by ϵ_M. The number of zeros of the solution $\sin\sqrt{M}(x+c)$ of the equation $y_{xx}+My=0$ in an interval of length ϵ_M is not less than the integral

part of $\frac{\sqrt{M}\epsilon_M}{\pi}$. It follows immediately from § 1, I that the number of zeros of solutions of (10) is not less than that of $\sin\sqrt{M}(x+c)$. Hence the

THEOREM II. *In order that there be N zeros in a solution of* (10) *it is sufficient that one can find an M such that* $\frac{\sqrt{M}\epsilon_M}{\pi}>N$, *where* ϵ_M *is the length of a sub-interval where* $G(x,\lambda)>MP$.

COROLLARY I. *In order that the number of zeros be unlimited it is sufficient that* $\lim\limits_{\lambda=\infty}\epsilon_M\sqrt{M}=\infty$.

COROLLARY II. *If throughout any sub-interval of fixed length the value of* $G(x,\lambda)$ *increases without limit, the number of zeros increases indefinitely.*

However, the sum of the sub-intervals in which $G>M$ may remain above a positive constant without compelling the number of zeros to increase with M. For example, one can set up a function which has no zeros except at $x=0$ and $x=1$ and which, except in the neighborhood of these points, oscillates between $y=\frac{1}{2}$ and $y=\frac{3}{4}$, the number of oscillations increasing indefinitely with λ. Moreover, by taking the oscillations frequent enough one can have $G(x,\lambda)>MP$ in portions which total at least one-quarter (or any other proper fraction) of the interval, the great curvature downward in these sub-intervals being counterbalanced by curvature upward in those where $G(x,\lambda)$ has the opposite sign.

On the other hand, we have seen that the number of zeros of the solution $\sin\sqrt{M}(x+c)$ of the equation $y_{xx}+My=0$ is not less than the integral part of $\frac{\sqrt{M}\epsilon_M}{\pi}$ where ϵ_M is the length of the interval. If this interval is divided into two parts ϵ_1, ϵ_2 for which there are solutions $\sin\sqrt{M}(x+c_1)$, $\sin\sqrt{M}(x+c_2)$ respectively, the number of zeros can not be reduced by more than one. For, the sum of the integral parts of $\frac{\sqrt{M}\epsilon_1}{\pi}$ and $\frac{\sqrt{M}\epsilon_2}{\pi}$ cannot differ by more than one from the integral part of $\frac{\sqrt{M}(\epsilon_1+\epsilon_2)}{\pi}$. It follows in the same way that if the interval is divided into n_1 parts the number of zeros of the solutions can not be decreased by more than n_1-1. But in the intervals where $G(x,\lambda)>MP$ the zeros of the solutions of (10) must be at least as many as the minimum number for $y_{xx}+My=0$.

THEOREM III. *In order that the solution of* (10) *have N zeros it is sufficient that one can find an M such that* $\frac{\sqrt{M}\eta_M}{\pi}-n_1+1>N$, *where* η_M *denotes the sum of the lengths of the* n_1 *intervals in which* $\frac{G}{\max p(x)}>M$.

§ 3. *Behavior of the Zeros with Monotone Change of λ.*

Let us fix the zero at the left-hand end of the interval 0, 1, denote such a solution by Y $[Y(0)=0]$, and investigate what happens to the other zeros as λ increases. It is convenient to think of $G(x, \lambda)$ as being defined for values $x>1$. Since the coefficients of (10) are continuous, the zeros move continuously, and since we can not have both $y(x)=0$ and $y'(x)=0$ without having $y\equiv 0$, the zeros can not coalesce and then disappear. By differentiating (10) with regard to λ we have

$$\frac{\partial(py_x)_x}{\partial\lambda}+G\frac{\partial y}{\partial\lambda}+\frac{\partial G}{\partial\lambda}y=0, \tag{12}$$

and on multiplication of this equation by $-y$ and of (10) by $\frac{\partial y}{\partial\lambda}$, addition and integration from x_1 to x_2 we get the fundamental formula*

$$\left[py_x\frac{\partial y}{\partial\lambda}\right]_{x_1}^{x_2}-\left[py\frac{\partial y_x}{\partial\lambda}\right]_{x_1}^{x_2}=\int_{x_1}^{x_2}\frac{\partial G}{\partial\lambda}y^2dx. \tag{13}$$

For the particular solution $Y(x)$ we have $Y(0)=0, \frac{\partial Y(0)}{\partial\lambda}$, and if α is another zero of Y the formula becomes

$$p(\alpha)Y_x(\alpha)\frac{\partial Y}{\partial\lambda}(\alpha)=\int_0^\alpha\frac{\partial G}{\partial\lambda}Y^2dx. \tag{13'}$$

If the integral on the right is positive, the signs of $Y_x(\alpha)$ and $\frac{\partial Y(\alpha)}{\partial\lambda}$ are the same. Hence when $Y_x(\alpha)$ is negative, $Y(\alpha)$ is decreasing with increase of λ, and this zero of Y is moving to the left; when $Y_x(\alpha)$ is positive, $Y(\alpha)$ is increasing and the zero is again moving to the left. In the same way it may it may be shown that all zeros for which the integral on the right of (13′) is negative, move to the right. When the integral is zero, further investigation is needed. We have now proved

---

*If p is also a function of λ this formula becomes

$$\left[py_x\frac{\partial y}{\partial\lambda}\right]_{x_1}^{x_2}-\left[py\frac{\partial y_x}{\partial\lambda}\right]_{x_1}^{x_2}-\left[\frac{\partial p}{\partial\lambda}y_xy\right]_{x_1}^{x_2}=\int_{x_1}^{x_2}\frac{\partial G}{\partial\lambda}y^2dx-\int_{x_1}^{x_2}\frac{\partial p}{\partial\lambda}y^2_xdx.$$

In terms of this hypothesis results analogous to all the succeeding theorems of this paper may be written down. Since, however, by a transformation of variables p can be made independent of λ we shall, for the sake of simplicity, confine ourselves to the problem proposed.

If at $x=0$ the condition $y_x(0)+\sigma y(0)=0$ is imposed where σ is a constant, then $\frac{\partial y_x(0)}{\partial\lambda}+\sigma\frac{\partial y(0)}{\partial\lambda}=0$, and this formula (13) can be written

$$-y^2p\frac{d}{dx}\left(\frac{\frac{\partial y}{\partial\lambda}}{y}\right)\equiv p\left[y_x\frac{\partial y}{\partial\lambda}-y\frac{\partial y_x}{\partial\lambda}\right]=\int_0^x\frac{\partial G}{\partial\lambda}y^2dx.$$

Hence, in order that the roots of a solution y and the roots of $\frac{\partial y}{\partial\lambda}$ alternate it is sufficient that $\int_0^x\frac{\partial G}{\partial\lambda}y^2dx$ have one sign.

THEOREM IV. *If $Y(0)=0$ and if λ is a parameter value for which there are subsequent zeros of Y at $\alpha_1, \ldots, \alpha_n$ then with increase of λ a zero, α_i, moves to the left or right according as*

$$\int_0^{\alpha_i} \frac{\partial G(x,\lambda)}{\partial \lambda} Y^2(x,\lambda)\,dx$$

is positive or negative.

The argument used in deducing *Theorem* IV is still valid if in place of a zero of y we introduce a zero of y_x.

THEOREM IV A. *If $y(0)=0$ or $y_x(0)=0$ and if λ is a parameter value such that $y(\alpha)=0$ or $y_x(\alpha)=0$, then with increase of λ the zero α moves to the left or right according as*

$$\int_0^{\alpha} \frac{\partial G(x,\lambda)}{\partial \lambda} y^2(x,\lambda)\,dx$$

is positive or negative.

COROLLARY. *In order that λ be a multiple characteristic number and $Y(x,\lambda)$ a multiple solution of the problem* (10) (6) *it is necessary** *that*

$$\int_0^1 \frac{\partial G(x,\lambda)}{\partial \lambda} Y^2(x,\lambda)\,dx=0.$$

G can be so chosen as a function of λ that with increasing λ the integral $\int_0^1 \frac{\partial G}{\partial \lambda} Y^2 dx$ is alternately positive and negative, and since the corresponding intervals of λ may be small at pleasure, we see that a zero may pass back and forth through $x=1$ an infinite number of times. There may then be an infinite number of λ's for which exist solutions of (10) (6) with a fixed number of zeros.

As λ runs through its interval there will be at least one value for which the number of zeros of $Y(x,\lambda)$ is a minimum. For both the orthogonal and polar cases this minimum number of oscillations is zero, but in general this will not be the case. If for $\lambda=\lambda'$ there is a minimum number n_1 of oscillations, then from the principle of continuity we can argue that for $n>n_1$ there is under *Hypothesis A* (§ 2) at least one value of λ for which exists a solution of (10) vanishing at $x=0$ and $x=1$, and oscillating n times (having n zeros) in the interval, while under *Hypothesis B* there are at least two.

One fundamental problem for investigation is the determination of sufficient conditions that in a given interval of λ (which may include infinity) there is *only one* solution of (10) (6) oscillating a given number of times; in

* This condition may be shown to be sufficient also.

other words, sufficient conditions that the zeros of $Y(x)$ pass through the point $x=1$ in one direction only as λ increases (or decreases). Another problem is that of determining when there exist precisely two solutions oscillating a given number of times.

THEOREM V. *If for values of λ in an interval (which may include infinity) there are solutions of* (10) (6) *for which the minimum and maximum number of oscillations are n_1 and n_2 (n_2 may be infinite), respectively, and if the value of $\frac{\partial G}{\partial \lambda}$ is equal to or greater than $\phi(\lambda)G(x,\lambda)$, where ϕ is a positive function of λ, then there is one and only one solution oscillating n times ($n_1 \leqq n \leqq n_2$) in the interval $0 \leqq x \leqq 1$.*

For, we have on multiplying (10) by Y, and integrating under the boundary conditions (6),

$$\int_0^1 pY_x^2 dx = \int_0^1 GY^2 dx. \tag{14}$$

From the hypothesis we may then write

$$\int_0^1 \frac{\partial G}{\partial \lambda} Y^2 dx \geqq \phi(\lambda) \int_0^1 GY^2 dx = \phi(\lambda) \int_0^1 pY_x^2 dx > 0$$

and since this holds for all λ, *Theorem* IV gives the desired result.

The theorem just proved fits the case where with unlimited increase of λ the maximum of G increases without limit. For the case where λ approaches $-\infty$, we observe from *Theorem* IV that a sufficient condition for a similar theorem is that $\int_0^{\prime} \frac{\partial G}{\partial \lambda} Y^2 dx$ be negative. To insure this, it is sufficient to take $\frac{\partial G}{\partial \lambda} < -\psi G$, where $\psi(\lambda)$ is a positive function. Then

$$\int \frac{\partial G}{\partial \lambda} Y^2 dx < -\psi(\lambda) \int GY^2 dx = -\psi(\lambda) \int pY_x^2 dx < 0.$$

Combining this result with *Theorem* V we may state the following:

THEOREM VI. *If for two values λ', λ'' ($\lambda' \geqq \lambda''$) of λ there are n_1 zeros of the solution Y within the interval* 0, 1, *and if for $\lambda > \lambda'$, $\frac{\partial G}{\partial \lambda} \geqq \phi_1(\lambda) G(x,\lambda)$, and for $\lambda < \lambda''$, $\frac{\partial G}{\partial \lambda} \leqq -\phi_2(\lambda) G(x,\lambda)$ where ϕ_1, ϕ_2 are positive functions defined in the intervals λ', $+\infty$ and $-\infty$, λ'', respectively, then under Hypothesis B* (§ 2) *for $n \geqq n_1$ there are two and only two values of λ for which there is a solution $Y(x,\lambda)$ of* (10) (6) *with n zeros.*

In the polar problem we have the special case of the hypothesis of this theorem where $\lambda'=\lambda''=0$, $n_1=0$, $G=q+\lambda k$, $\frac{\partial G}{\partial \lambda}=k$, $\phi_1=\frac{1}{\lambda}>0$, $-\phi_2=\frac{1}{\lambda}<0$. A more general special case of *Theorem* VI is discussed in § 4.

Since formula (14) is valid also when

$$y(0)=y_x(1)=0, \text{ or } y_x(0)=y(1)=0, \text{ or } y_x(0)=y_x(1)=0,$$

theorems analogous to VI can be written down for any of these boundary conditions.

§ 4. *The Non-Definite Case.*

With regard to applications, the differential equation

$$(py_x)_x+(q+\lambda k)y=0, \tag{15}$$

with boundary conditions

$$y(0)=y(1)=0, \tag{16}$$

is the most important special type of (10). The problem where one of the integrals

$$D(y, 0, 1)=\int_0^1 (py_x^2-qy^2)\,dx, \quad \int_0^1 ky^2dx$$

is definite, has been studied in great detail. Let us first investigate some properties of the remaining *non-definite case* where both these integrals may take on negative values and later deduce some analogous results for the more general equation (10). In §§ 6–8 these results will be extended to cover the more general boundary condition (2).

For $q\leqq 0$ the integral $D(y, 0, 1)$ can not be negative. If q is positive in at least a part of the interval 0, 1, the minimum of the integral $\int_0^1 py_x^2dx$ under the conditions (16) and $\int_0^1 qy^2dx=1$ is (§ 1, III) the smallest parameter value λ of the equation $(py_x)_x+\lambda qy=0$, and when $\frac{q}{\min p}\leqq\pi^2$, this in turn is (§ 1, I) at least as great as the smallest of the equation $y_{xx}+\lambda\pi^2y=0$ under the same boundary conditions (16). Since for the latter equation $\lambda_1=1$ and $Y_1=c\sin\pi x$, it follows that λ is greater than unity. Hence $\int_0^1 py_x^2dx=\lambda$ is greater than $\int_0^1 qy^2dx=1$ and $D(y, 0, 1)$ is positive.

In the non-definite case one must therefore have $\frac{q}{\min p}>\pi^2$ *in at least a part of the interval,* and this assumption we shall now make. But we note here that in a sub-interval α_1, α_2 such that $y(\alpha_1)=y(\alpha_2)=0$, any value of q for which $\frac{q}{\min p}<\frac{\pi^2}{(\alpha_2-\alpha_1)^2}$, makes the integral $\int_{\alpha_1}^{\alpha_2}(py_x^2-qy^2)\,dx$ definite, as is

readily shown by a comparison argument like that above. Hence, for given p and q, if throughout the interval 0, 1 the zeros can be taken close enough together (as *e.g.* may certainly be done where λ has the same sign as $k(x)$ and is taken large enough in absolute value, cf. § 1, I), the integral for each subinterval is definite, and hence $D(y, 0, 1)$ is positive. *This is also true under the more general boundary conditions* (2).

On the other hand from the formula

$$D(y, \alpha_1, \alpha_2) \equiv \int_{\alpha_1}^{\alpha_2} (py_x^2 - qy^2)\,dx = \lambda \int_{\alpha_1}^{\alpha_2} ky^2\,dx, \tag{17}$$

which is obtained by multiplying (15) by y and integrating under the conditions $y(\alpha_1) = y(\alpha_2) = 0$, we see that if λ is positive and k negative in the interval, then $D(y, \alpha_1, \alpha_2)$ is negative; it is, however, still possible that for the larger interval 0, 1 the integral $D(y, 0, 1)$ be positive. Were this the case (as will later be established provided λ is large enough) we could deduce from *Theorem IV* and the formula

$$D(y, 0, 1) \equiv \int_0^1 (py_x^2 - qy^2)\,dx = \lambda \int_0^1 ky^2\,dx, \quad y(0) = y(1) = 0, \tag{18}$$

certain facts in regard to the movement of the subsequent zeros of particular solutions Y of (15) which vanish at $x = 0$. Before proceeding to a detailed discussion of this matter let us compare some possibilities for the orthogonal, polar and non-definite cases.

In the orthogonal case let us consider all λ greater than some finite number chosen less than the smallest characteristic (which may be positive or negative). As λ increases, the subsequent zeros of any solution Y of (15) $[Y(0) = 0]$ move to the left, new ones being added to the interval 0, 1. For the smallest value of λ there are no zeros within the interval and when λ passes through a parameter value λ_n the n-th zero enters. In the polar case there are two ranges of values of λ; for the range which extends from zero to $+\infty$, a result precisely like that of the orthogonal case may be stated; for the range which extends from $-\infty$ to zero, a decrease of λ causes subsequent zeros to move to the left, there being no zero of $Y(x, 0)$ present in the interval while the n-th enters for $\lambda = \lambda_{-n}$. In the non-definite case we know that the results must be quite different and we may expect that the march of the zeros will not be monotone with λ. In fact, *there may be a range of values of* λ $(L_1 \leqq \lambda \leqq L_2)$ *such that as* λ *increases the number of zeros first decreases, then increases, then decreases and finally increases, the minimum number being a positive integer.**

* A reference to the proof of this last fact is given in the Introduction.

As an example let us consider the equation

$$y_{xx}+[(100\pi)^2+\lambda\varkappa)y=0, \quad y(0)=y(1)=0,$$

where $\varkappa$ is equal -1 in the sub-interval $0, \frac{1}{2}$ and equal 1 in the remainder of the interval. Such a function $\varkappa(x)$ may be approximated by an analytic function $k(x)$ for which the corresponding equation has solutions with similar properties. When $\lambda=0$ the solution $y=\sin 100\pi x$ has 100 zeros in the interval. When $\lambda=(100\pi)^2$, we have the equation $y_{xx}=0$ in $0, \frac{1}{2}$ and $y_{xx}+2(100\pi)^2y=0$ elsewhere, the solutions being respectively $y=cx$ (c=const.), $y=\sin 100\sqrt{2}\pi x$. There are are then approximately $50\sqrt{2}$ zeros. For $\lambda=-(100\pi)^2$ there are evidently the same number of zeros. More generally, when λ is in the interval $-(100\pi)^2$, $(100\pi)^2$ the number of zeros is approximately

$$\frac{1}{2\pi}[\sqrt{(100\pi)^2+\lambda}+\sqrt{(100\pi)^2-\lambda}],$$

a function which has its maximum for $\lambda=0$ and *decreases* with increase of $|\lambda|$. On the other hand when $|\lambda|$ is greater than $(100\pi)^2$ the number of zeros is approximately $\frac{1}{2\pi}\sqrt{(100\pi)^2+|\lambda|}$, which *increases* with $|\lambda|$. The minimum number of zeros is then approximately $50\sqrt{2}$ and occurs for $\lambda=(100\pi)^2$ and $-(100\pi)^2$.* *For values of n between* $50\sqrt{2}$ *and* 100 *there are four parameters corresponding, to which exist solutions oscillating n times.*

Returning now to the discussion of the sign of $D(y, 0, 1)$ we can prove

LEMMA I. *If in an interval a, b, the function k is positive, except perhaps at the end points, then λ can be taken so large that for solutions of* (15),

$$D(y, a, b)\equiv\int_a^b(py_x^2-qy^2)\,dx>0.$$

For, by taking λ large enough, $q+\lambda k$ can be made as large as desired, except perhaps at the end points. Hence the zeros are as thickly strewn throughout the interval a, b as desired (§ 1, I). We note that in any subinterval β_1, β_2 of a, b, in which there is a zero of y and for which the maximum of $|y|$ is M, the minimum value of $\int_{\beta_1}^{\beta_2}y_x^2dx$ is taken on when y is a straight line joining $(\beta_1, 0)$ and $(\beta_2, \pm M)$. This minimum value is $\frac{M^2}{\beta_2-\beta_1}$. Since on the other hand $\int_{\beta_1}^{\beta_2}qy^2dx<(\beta_2-\beta_1)M^2\max q$, the value of $D(y, \beta_1, \beta_2)$ is greater than $M^2\left[\frac{\min p}{\beta_2-\beta_1}-(\beta_2-\beta_1)\max q\right]$: hence when $(\beta_2-\beta_1)^2$ is made less than

* By reference to *Theorem* IV it follows from this discussion that in the two intervals which are approximately $-(100\pi)^2, 0$; $(100\pi)^2, +\infty$ the value of $\int_0^1\varkappa y^2dx$ must be positive or $\int_0^{\frac{1}{2}}y^2dx<\int_{\frac{1}{2}}^1 y^2dx$, while in the intervals which are approximately $-\infty, -(100\pi)^2$; $0, (100\pi)^2$, $\int_0^{\frac{1}{2}}y^2dx>\int^1 y^2dx$.

$\frac{\min p}{\max q}$ this integral is definite. It follows that since the first and last zeros of $y(x)$ are as close to the ends of the interval a, b as is desired, those portions of the integral arising from these two end sub-intervals can be made positive. For any interval α_1, α_2 of oscillation $[y(\alpha_1)=y(\alpha_2)=0]$ it follows at once from (17) that when k and λ are positive $D(y, \alpha_1, \alpha_2)>0$. Combining these results we have the lemma.

It follows in the same way that if $k \leqq 0$ and $k \not\equiv 0$ in an interval a, b, then λ can be taken so large negatively that for solutions of (15) the integral $D(y, a, b)$ is positive.

Coming now to an interval c, d in which k is negative, except perhaps at c and d, it is possible in an interval $c+\eta$, $d-\eta$ ($\eta>0$ arbitrarily small) to take M' arbitrarily large and then choose λ so large that $-(q+\lambda k)>M'$ or

$$|(py_x)_x| = |-(q+\lambda k)y| > M'|y|, \quad [(py_x)_x > M'y \text{ if } y>0]. \qquad (19)$$

The arc $y(x)$ may be shown to be sharply concave away from the x-axis.* It is possible to prove the following:

Lemma II. *If in an interval c, d the function k is negative, except perhaps at the end points, then λ can be taken so large that for a solution of* (15)

$$D(y, c, d) \equiv \int_c^d (py_x^2 - qy^2)\,dx > 0.$$

This will be proved by showing that on taking λ great enough it is possible to insure that $\frac{y_x^2}{y^2}$ is great at pleasure except perhaps in sub-intervals whose length decreases indefinitely with $\frac{1}{\lambda}$. Since in the interval $c+\eta$, $d-\eta$ the curve is concave upward for positive y and concave downward for negative y, there can not be more than one zero. The discussion may be separated into two parts according as $y(x)$ has a zero or not.

If $y(\gamma)=0$, then without loss of generality it may be assumed that within the interval γ, $d-\eta$, y is positive and hence, since $y_{xx}>0$, that y_x is positive. Multiplying the inequality (19) by $2py_x$ and integrating we get

$$\int_\gamma^x 2py_x(py_x)_x dx > M' \int_\gamma^x 2ypy_x dx,$$

and hence

$$p^2(x)y_x^2(x) \geqq p^2(x)y_x^2(x) - p^2(\gamma)y_x^2(\gamma) \geqq M'[\min p] \int_\gamma^x 2yy_x dx$$
$$= M'y^2(x) \min p, \quad (20)$$

---

* This is at once evident if the equation is taken in the form $\bar{y}_{xx} + (\bar{q} + \lambda\bar{k})\bar{y} = 0$ to which it may be reduced by a transformation.

for all x in the interval γ, $d-\eta$. Similar reasoning establishes the same result in the interval $c+\eta$, γ. Hence, for the case that y vanishes, we have throughout the interval $c+\eta$, $d-\eta$ the inequality $\frac{y_x^2}{y^2} > \frac{M' \min p}{p^2}$, which as noted above, is sufficient to establish the theorem.

If on the other hand y has no zero in the interval, let us take the function positive and denote by γ the point at which it takes on its minimum. It may be assumed that γ is not at the right-hand end of the interval; that special case could be treated in an analogous fashion. By multiplying y by a constant, $y(\gamma)$ can be made equal 1, while the sign of the integral $D(y, c, d)$ is not altered. Since $y_x(\gamma) \geqq 0$, we have $y>1$ in the neighborhood of γ; in fact, for any arbitrarily small but fixed η we can, by taking M' large enough, have $y(\gamma+\eta)>2$. For, since $y \geqq 1$ it follows from the inequality (19) that

$$py_x \geqq py_x - py_x(\gamma) = \int_\gamma^x (py_x)_x dx \geqq M' \int_\gamma^x y dx = M'(x-\gamma).$$

Integrating again,

$$(y-1) \max p = \max p \int_\gamma^x y_x dx \geqq \int_\gamma^x py_x dx \geqq M' \int_\gamma^x (x-\gamma) dx = \frac{M'(x-\gamma)^2}{2}.$$

By taking $x-\gamma \geqq \eta$ and choosing $\frac{M'\eta^2}{2 \max p} > 1$ we have $y-1>1$ or $y>2$. From this it follows that on multiplying the inequality (19) by $2py_x$ and integrating we get by a process similar to that used in (20) the formula

$$p^2(x) y_x^2(x) > M' \int_\gamma^x 2ypy_x dx \geqq M'[y^2(x)-1] \min p > \frac{3M'}{4} y^2(x) \min p.$$

Hence in the interval $\gamma+\eta$, $d-\eta$ we have $\frac{y_x^2}{y^2} > \frac{3M'}{4} \frac{\min p}{p^2}$. And since in the interval $c+\eta$, $\gamma-\eta$ a similar inequality may be obtained, this completes the proof of the lemma.

It follows in the same way that if $k>0$, except perhaps at the end points of the interval c, d, then λ can be taken so large negatively that for solutions of (15) the integral $D(y, c, d)$ is positive.

We have then proved by these lemmas that in any sub-interval in which k has one sign, $D(y)$ can be made positive by taking λ large enough positively and also by taking λ large enough negatively. It follows that $D(y, 0, 1)$ would be positive, and from the formula

$$\lambda \int_0^1 ky^2 dx = \int_0^1 (py_x^2 - qy^2) dx \equiv D(y, 0, 1) > 0$$

and *Theorem* IV, that for $|\lambda|$ large enough the zeros of Y move to the left with increase of $|\lambda|$. We can then enunciate

THEOREM VII. *There exists an integer n_2 such that for $n \geqq n_2$ there are precisely two solutions of* (15) (16) *oscillating n times.*

If proper restrictions are imposed on $G(x, \lambda)$, a theorem similar to the preceding may be proved for the general equation

$$(py_x)_x + G(x, \lambda) y = 0. \tag{10}$$

The proof used in *Theorems* VI and VII, and *Lemmas* I and II is valid also under the following hypotheses on a function $G'(x, \lambda)$ obtained by subtracting from $G(x, \lambda)$ a function $q(x) \cdot [G'(x, \lambda) \equiv G(x, \lambda) - q(x)]$.

(α) In one or more sub-intervals of 0, 1, $G'(x, \lambda)$ is a monotone increasing function of λ such that $\lim\limits_{\lambda=\infty} G'(x, \lambda) \geqq 0$ and for at least a part of each sub-interval $\lim\limits_{\lambda=\infty} G(x, \lambda) = +\infty$.

(β) In the remaining sub-interval or sub-intervals of 0, 1, $G'(x, \lambda)$ is a decreasing function of λ such that $\lim\limits_{\lambda=-\infty} G'(x, \lambda) \leqq 0$, and for at least a part of each sub-interval $\lim\limits_{\lambda=-\infty} G'(x, \lambda) = -\infty$.

(γ) For values λ', λ'' of λ there are n_1 zeros of the solution Y within the interval 0, 1 and for $\lambda > \lambda'$,

$$\frac{\partial G'}{\partial \lambda} \geqq \phi_1(\lambda) G'(x, \lambda), \text{ and for } \lambda < \lambda'', \frac{\partial G'}{\partial \lambda} \leqq -\phi_2(\lambda) G'(x, \lambda),$$

where ϕ_1, ϕ_2 are positive functions defined in the intervals λ', $+\infty$ and $-\infty$, λ'' respectively.

THEOREM VIII. *Under the conditions* (α) (β) (γ) *there exists an integer n_2 such that for $n > n_2$ there are precisely two solutions of* (10) (16) *oscillating n times.*

When the parameter value and corresponding solution of (15) (16) are complex, let us set $\lambda \equiv \sigma + i\tau$, $y \equiv u + iv$. The differential equation resolves itself into the two following:

$$(pu_x)_x + qu + \sigma ku - \tau kv = 0, \quad u(0) = u(1) = 0, \tag{21}$$

$$(pv_x)_x + qv + \sigma kv + \tau ku = 0, \quad v(0) = v(1) = 0. \tag{22}$$

Corresponding to $\lambda \equiv \sigma - i\tau$ there is then a solution $y \equiv u - iv$. On multiplying (21) and (22) in the first place by u and v, and in the second by v and $-u$ respectively, adding and integrating, one obtains the two formulae

$$\int_0^1 [p(u_x^2 + v_x^2) - q(u^2 + v^2)] dx = 0, \quad \int_0^1 k(u^2 + v^2) dx = 0. \tag{23}$$

More generally one obtains by this process the formula

$$[p(u_x v - v_x u)]_{x=\alpha}^{x=\beta} = \tau \int_\alpha^\beta k(u^2+v^2)\,dx. \tag{24}$$

The orthogonality of two solutions ($\int_0^1 kY_n Y_m dx = 0$, $m \neq n$) holds as well for complex as for real characteristic numbers λ_n, λ_m. Hence by separation into real and imaginary parts we get the

THEOREM IX. *If* $Y_m(x, \lambda_m) \equiv u_m + iv_m$, $Y_n \equiv u_n + iv_n$ *are two solutions of* (15) (16), *then*

$$\int_0^1 k(u_m u_n - v_m v_n)\,dx = 0, \quad \int_0^1 k(u_m v_n + u_n v_m)\,dx = 0.$$

Theorems of other types may be derived of which the following is an example:

THEOREM X. *If in the interval* 0, 1, *the function* $k(x)$ *changes sign once only, then the roots of the real and imaginary parts* u_n, v_n *of the solution* $Y_n \equiv u_n + iv_n$ *of* (15) (16) *separate one another.*

For, formula (24) may be written

$$p(u_x v - v_x u) = \int_0^x k(u^2+v^2)\,dx, \text{ or } p\,\frac{d\left(\frac{u}{v}\right)}{dx} = \frac{1}{v^2}\int_0^x k(u^2+v^2)\,dx.$$

From the hypothesis that k changes sign but once it follows that the integral can not vanish within the interval, and since $\dfrac{d\left(\frac{u}{v}\right)}{dx} > 0$ except at the end points, $\dfrac{u}{v}$ is a monotone function and the theorem follows at once. We may note further that under the hypotheses of the theorem neither u nor v can gain or lose a zero at a point within the interval. For, were there a zero lost or gained at $x = a$, both $y(a)$ and $y_x(a)$ would vanish, and formula (24) could be written $\int_0^a k(u^2+v^2)\,dx = 0$, which would give a contradiction.

§ 5. *Reduction of the General Boundary Conditions to Normal Forms.*

In the preceding sections the discussion has dealt mainly with the simple boundary conditions $y(0)y_x(0) = y(1)y_x(1) = 0$. To facilitate the discussion of the most general boundary conditions it is desirable to obtain normal forms.* The linear self-adjoint equation of the second order will be taken in the form

$$(\pi(x)u_x(x))_x + \Gamma(x, \lambda)u(x) = 0, \quad \pi(x) > 0, \tag{25}$$

* The classification used in this section follows that of Hilbert and Haupt, *loc. cit.* The geometrical form into which the transformation is thrown was suggested by my colleague, Prof. H. P. Manning.

and if the linearly independent boundary conditions

$$\left.\begin{aligned}\alpha_1 u(0)+\alpha_2 u_x(0)+\alpha_3 u(1)+\alpha_4 u_x(1)=0,\\ \beta_1 u(0)+\beta_2 u_x(0)+\beta_3 u(1)+\beta_4 u_x(1)=0,\end{aligned}\right\} \tag{26}$$

are to be self-adjoint, the condition

$$\pi(0)\begin{vmatrix}\bar{u}(0) & \bar{u}_x(0)\\ \bar{\bar{u}}(0) & \bar{\bar{u}}_x(0)\end{vmatrix}=\pi(1)\begin{vmatrix}\bar{u}(1) & \bar{u}_x(1)\\ \bar{\bar{u}}(1) & \bar{\bar{u}}_x(1)\end{vmatrix} \tag{27}$$

must be imposed, where $\bar{u}$, $\bar{\bar{u}}$ are any functions with continuous derivatives which satisfy (26). We have the relation

$$\begin{vmatrix}\alpha_1 & \alpha_2\\ \beta_1 & \beta_2\end{vmatrix}\cdot\begin{vmatrix}\bar{u}(0) & \bar{u}_x(0)\\ \bar{\bar{u}}(0) & \bar{\bar{u}}_x(0)\end{vmatrix}=\begin{vmatrix}\alpha_3 & \alpha_4\\ \beta_3 & \beta_4\end{vmatrix}\cdot\begin{vmatrix}\bar{u}(1) & \bar{u}_x(1)\\ \bar{\bar{u}}(1) & \bar{\bar{u}}_x(1)\end{vmatrix}, \tag{28}$$

as may be seen by applying the usual rule for multiplying determinants and noting from (26) that each of the four elements of one product determinant is the negative of the corresponding element of the other. Let us denote by B_{ij} the determinant formed by taking the i-th and j-th columns of the matrix

$$\begin{Vmatrix}\alpha_1 & \alpha_2 & \alpha_3 & \alpha_4\\ \beta_1 & \beta_2 & \beta_3 & \beta_4\end{Vmatrix}.$$

Considering the two determinants of the u's as the variables, we have from the theory of linear equations that a necessary and sufficient condition for the solution of (27) and (28) is that

$$\pi(1)B_{12}-\pi(0)B_{34}=0. \tag{29}$$

Hence B_{12} and B_{34} are simultaneously zero or different from zero.

Let us subject the dependent variable to the transformation

$$u(x)=\eta(x)y(x),\quad \eta(x)\neq 0, \tag{30}$$

by which the number of zeros of the solution remains unaltered. The equation (25) takes on the self-adjoint form

$$(py_x)_x+G(x,\lambda)y=0,\quad p=\pi\eta^2>0,\quad G(x,\lambda)=\pi\eta\eta_{xx}+\pi_x\eta\eta_x+\Gamma\eta^2, \tag{31}$$

and the conditions (26) are replaced by a similar set

$$\left.\begin{aligned}\gamma_1 y(0)+\gamma_2 y_x(0)+\gamma_3 y(1)+\gamma_4 y_x(1)=0,\\ \delta_1 y(0)+\delta_2 y_x(0)+\delta_3 y(1)+\delta_4 y_x(1)=0.\end{aligned}\right\} \tag{32}$$

We have at once from (31) the important formula

$$\int_0^1\frac{\partial\Gamma}{\partial\lambda}u^2dx=\int_0^1\frac{\partial G}{\partial\lambda}y^2dx. \tag{33}$$

Writing u_1, u_2, u_3, u_4, for $u(0), u_x(0), u(1), u_x(1)$, and using a similar notation for the η's and $y's$ we can interpret our problem in relation to the tetraedron of reference as the investigation of the line (26) subject to the condition (27) or (29). If $B_{34}=0$, the u_3 and u_4 can be eliminated from (26) at the same time, giving

$$B_{13}u_1+B_{23}u_2=0 \text{ or } B_{14}u_1-B_{42}u_2=0, \tag{34}$$

either of these equations representing the plane determined by the given line and the edge $12(u_1=0, u_2=0)$. Hence the line (26) intersects this edge of the tetraedron. But if $B_{34}=0$ we have from (29) $B_{12}=0$, and the straight line intersects also the edge 34. The equation of the plane determined by the given line and the edge 34 can be written

$$B_{13}u_3+B_{14}u_2=0 \text{ or } B_{23}u_3-B_{42}u_2=0. \tag{35}$$

From (34) and (35) we see that under the hypothesis $B_{12}=B_{34}=0$ the boundary conditions may be written

Case I. $$\sigma u_1+u_2=0, \quad \tau u_3+u_4=0, \tag{36}$$

where

$$\sigma=\frac{B_{13}}{B_{23}}=-\frac{B_{14}}{B_{42}}, \quad \tau=\frac{B_{13}}{B_{14}}=-\frac{B_{23}}{B_{42}}.$$

The parameters σ and τ may have zero or infinite values when the line (26) lies in a face of the tetraedron.

In general, when any two $B's$ with complementary indices are zero, the line intersects two opposite edges of the tetraedron and the conditions (26) may be reduced to a normal form similar to (36).

The transformation (30) takes the form

$$u_1=\eta_1 y_1, \quad u_2=\eta_2 y_1+\eta_1 y_2, \quad u_3=\eta_3 y_3, \quad u_4=\eta_4 y_3+\eta_3 y_4, \quad \eta_1\neq 0, \eta_2\neq 0, \tag{37}$$

and corresponds to a rotation of the face 2 of the tetraedron about the edge 12, and of the face 4 about the edge 34, leaving the faces 1 and 2 and the edges 12, 13 and 34 unchanged. If we denote by A_{ij} the determinants of the matrix of the coefficients γ_i, δ_i of (32) we can read their values from the identity

$$\begin{vmatrix} \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 \\ \delta_1 & \delta_2 & \delta_3 & \delta_4 \end{vmatrix} = \begin{vmatrix} \eta_1\alpha_1+\eta_2\alpha_2 & \eta_1\alpha_2 & \eta_3\alpha_3+\eta_4\alpha_4 & \eta_3\alpha_4 \\ \eta_1\beta_1+\eta_2\beta_2 & \eta_1\beta_2 & \eta_3\beta_3+\eta_4\beta_4 & \eta_3\beta_4 \end{vmatrix};$$

thus $A_{12}=\eta_1^2B_{12}$, $A_{34}=\eta_3^2B_{34}$, etc. As may be seen from these formulae and (31), we have corresponding to (27) or (29), the relation

$$p(1)A_{12}-p(0)A_{34}=0 \tag{38}$$

which, when added to (32), makes these boundary conditions self-adjoint.

The discussion of the boundary conditions under the transformation may be sub-divided as follows:

I. If the given line intersects the edges 12 and 34 ($B_{34}=B_{12}=0$), the same will be true after the transformation since $A_{34}=A_{12}=0$. To reduce the condition to the normal form (36) (which we shall in general use) the transformation is superfluous. This case is called the sturmian.

From the formulae (37) it follows at once that the conditions (36) can by proper choice of $\eta_1, \eta_2, \eta_3, \eta_4$ be reduced to one of the simpler forms

$$\text{(a) } y_1=y_3=0; \quad \text{(b) } y_1=y_4=0; \quad \text{(c) } y_2=y_3=0; \quad \text{(d) } y_2=y_4=0, \quad (39)$$

these corresponding geometrically to the cases where the given line coincides respectively with the edge 13, the new edge 14, the new edge 23 or the new edge 24.

II. If the line does not intersect the edges 12, 34, but does intersect 13, the transformation may be so determined that it will also intersect 24; *e. g.*, by making the plane 2 pass through the intersection of the given line with the plane 4. In this case since u_4 is unaltered we have $\eta_3=1$, $\eta_4=0$, and hence to equate $A_{13}\equiv\eta_3(\eta_1 B_{13}+\eta_2 B_{23})$ to zero we need only to choose $\frac{\eta_2}{\eta_1}=-\frac{B_{13}}{B_{23}}$. After the transformation we have $A_{13}=A_{24}=0$, and by elimination in (32) we get by aid of (38)

Case II. $$y_1=hy_3, \quad hp(0)y_2=p(1)y_4,$$

where

$$h=\frac{A_{23}}{A_{12}}=-\frac{A_{34}}{A_{14}}=-\frac{p(1)A_{12}}{p(0)A_{14}}=\frac{p(1)A_{23}}{p(0)A_{34}}.$$

III. If the line does not intersect any of the three edges 12, 34, 13, ($B_{34}\neq 0$, $B_{12}\neq 0$, $B_{24}\neq 0$) we can determine the transformation so that the face 2 of the tetraedron shall pass through the intersection of the line and the face 3, while the face 4 passes through the intersection of the line and the face 1. Thus, in the new tetraedron, the line intersects the edges 23 and 14, and we have the relations $A_{14}\equiv\eta_3(\eta_1 B_{14}+\eta_2 B_{24})=0$, $A_{23}\equiv\eta_1(\eta_3 B_{23}+\eta_4 B_{24})=0$, which determine the ratios $\frac{\eta_1}{\eta_2}$, $\frac{\eta_3}{\eta_4}$ desired in this case. The equations of the line can now be written

Case III. $$y_1=lp(1)y_4, \quad lp(0)y_2=-y_3,$$

where

$$l=-\frac{A_{42}}{p(1)A_{12}}=\frac{A_{34}}{p(1)A_{13}}=\frac{A_{12}}{p(0)A_{13}}=-\frac{A_{42}}{p(0)A_{34}}.$$

THEOREM XI. *By a change of dependent variable the self-adjoint equation* (25) *with self-adjoint boundary conditions* (26) (27) *may be reduced to an equation* (31) *which is again self-adjoint, and for which the self-adjoint boundary condition may be written in one of three forms characterized respectively by* $A_{12}=A_{34}=0$; $A_{23}=A_{14}=0$; $A_{13}=A_{42}=0$:

Case I. $\sigma y(0)+y_x(0)=0, \quad \tau y(1)+y_x(1)=0,$ σ, τ *constants* (*including* ∞),
Case II. $y(0)=hy(1), \quad hp(0)y_x(0)=p(1)y_x(1),$ $h=$*constant*,
Case III. $y(0)=lp(1)y_x(1), \quad lp(0)y_x(0)=-y(1),$ $l=$*constant*,

The number of zeros of the solutions of the equation remains unaltered under this transformation.

The value of the integral $\int_0^1 \frac{\partial \Gamma}{\partial \lambda} u^2 dx$ *is the same as that of the corresponding integral* $\int_0^1 \frac{\partial G}{\partial \lambda} y^2 dx$.

COROLLARY I. *In Case I the transformation may be so chosen that the boundary conditions can be written in one of the special forms*

$$\text{(a)}\ y(0)=y(1)=0; \quad \text{(b)}\ y(0)=y_x(1)=0;$$
$$\text{(c)}\ y_x(0)=y(1)=0; \quad \text{(d)}\ y_x(0)=y_x(1)=0.$$

§ 6. *Oscillation Theorems for the Sturmian Boundary Conditions.*

We propose now to study oscillation theorems for solutions of the differential equation under the first or sturmian case of the general boundary conditions (26) (27). As was shown in the preceding section this case, characterized by the relations $B_{12}=B_{34}=0$ ($A_{12}=A_{34}=0$), may be reduced to a study of the equation

$$(py_x)_x+G(x,\lambda)y=0, \tag{41}$$

under the boundary conditions

$$\sigma y(0)+y_x(0)=0, \quad \tau y(1)+y_x(1)=0, \tag{42}$$

where σ and τ are constants, the important special forms

$$\left.\begin{array}{ll} \text{(a)}\ y(0)=y(1)=0; & \text{(b)}\ y(0)=y_x(1)=0; \\ \text{(c)}\ y_x(0)=y(1)=0; & \text{(d)}\ y_x(0)=y_x(1)=0; \end{array}\right\} \tag{43}$$

being considered as included for the values 0 and ∞. It was further shown that the boundary conditions may always be reduced to one of these special forms.

We proceed now to a study of the nature of the dependence of λ on σ and τ. If $\bar{y}$ and $\bar{\bar{y}}$ are defined as the following fundamental solutions of (41),

$$\bar{y}(0,\lambda)=0, \bar{y}_x(0,\lambda)=1; \quad \bar{\bar{y}}(0,\lambda)=1, \bar{\bar{y}}_x(0,\lambda)=0, \tag{44}$$

any solution $y(x)$ can be written $y=c_1\bar{y}+c_2$. For determination of c_1 and c_2 we have on substitution in (42),

$$\sigma c_2+c_1=0, \quad \tau[c_1\bar{y}(1,\lambda)+c_2\bar{\bar{y}}(1,\lambda)]+c_1\bar{y}_x(1,\lambda)+c_2\bar{\bar{y}}_x(1,\lambda)=0.$$

A necessary and sufficient condition that there be values of c_1, c_2 satisfying the equations is that

$$\begin{vmatrix} 1 & \sigma \\ \tau\bar{y}(1,\lambda)+\bar{y}_x(1,\lambda) & \tau\bar{\bar{y}}(1,\lambda)+\bar{\bar{y}}_x(1,\lambda) \end{vmatrix}=0.$$

Hence each of the parameters σ, τ is a one-valued function of the other and of λ. If we assume that a zero of the solution moves continuously to the right or left, and note what happens as two consecutive zeros pass through the point $x=0$, we see geometrically that for any given σ the condition $\sigma y(0)+y_x(0)=0$ must be satisfied at some stage of the process. It is possible by analysis to ascertain precisely what happens when σ or τ passes over the range $-\infty$, $+\infty$. Corresponding to $\sigma=-\infty$, 0, $+\infty$ one has respectively $y(0)=0$, $y_x(0)=0$, $y(0)=0$, and we shall show that under proper conditions one and only one zero of $y(x)$ is lost or gained by the process. A similar change in τ has a corresponding result.

If in formula (13) we substitute from (42) and the formulae

$$\sigma\frac{\partial y(0)}{\partial\lambda}+\frac{d\sigma}{d\lambda}y(0)+\frac{\partial y_x(0)}{\partial\lambda}=0, \quad \tau\frac{\partial y(1)}{\partial\lambda}+\frac{d\tau}{d\lambda}y(1)+\frac{\partial y_x(1)}{\partial\lambda}=0,$$

obtained from (42) by differentiation with regard to λ, we obtain the relation

$$\int_0^1\frac{\partial G}{\partial\lambda}y^2dx=-p(0)\frac{d\sigma}{d\lambda}y^2(0)+p(1)\frac{d\tau}{d\lambda}y^2(1). \tag{45}$$

Either of the parameters σ, τ may be held fixed. If τ is fixed $\frac{d\tau}{d\lambda}=0$, and in order that λ and σ be one-valued functions of one another over the range $-\infty$, $+\infty$ of σ it is then necessary and sufficient that the integral $\int_0^1\frac{\partial G}{\partial\lambda}y^2dx$ have one sign. From such considerations one may deduce a variety of theorems of which the following are examples.

THEOREM XIII. *If λ'_m, λ'_{m+1} denote two successive characteristic numbers of a solution of* (41) *for the boundary conditions* 43(*a*), *then in order that there be for all σ an intermediate value of λ corresponding to which there is precisely one solution of* (41) *for the boundary conditions $\sigma y(0)+y_x(0)=0$, $y(1)=0$, it is necessary and sufficient that the integral $\int_0^1\frac{\partial G}{\partial\lambda}y^2dx$ have one sign for all solutions $y(x)$ concerned.*

THEOREM XIV. *If $\bar{\lambda}_m$, $\bar{\lambda}_{m+1}$ denote two successive characteristic numbers of a solution of* (41) *for the boundary conditions*

$$\sigma y(0)+y_x(0)=0, \quad y(1)=0, \quad \sigma=constant,$$

then in order that there be for all τ an intermediate value of λ corresponding to which there is precisely one solution of (41) (42), *it is necessary and sufficient that the integral* $\int_0^1 \frac{\partial G}{\partial\lambda} y^2 dx$ *have one sign for all solutions* $y(x)$ *concerned.*

THEOREM XV. *If corresponding to λ', λ'' there are solutions of* (41) (43a) *with n_1 and n_2 zeros respectively, and if* $\int_0^1 \frac{\partial G}{\partial\lambda} y^2 dx$ *has one sign, then in the interval λ', λ'' there is one and only one value of λ for which* (41) (42) *has a solution $y(x)$ with n zeros* $(n_1<n<n_2)$.

We have seen in § 5 that by a transformation of the dependent variable the equation (41) may be thrown into the form

$$(\bar{p}\bar{y}_x)_x+\bar{G}\bar{y}=0, \quad \int_{\alpha_1}^{\alpha_2} \frac{\partial \bar{G}}{\partial\lambda} \bar{y}^2 dx=\int_{\alpha_1}^{\alpha_2} \frac{\partial G}{\partial\lambda} y^2 dx, \quad \bar{p}(x)>0,$$

where the boundary conditions (42) assume one of the special forms

$\bar{y}(\alpha_1)=\bar{y}(\alpha_2)=0$; $\bar{y}(\alpha_1)=\bar{y}_x(\alpha_2)=0$; $\bar{y}_x(\alpha_1)=\bar{y}(\alpha_2)=0$; $\bar{y}_x(\alpha_1)=\bar{y}_x(\alpha_2)=0$.

As an extension of *Theorems* IV and IVA we have then the following:

THEOREM XVI. *If a zero of one of the functions $\sigma y+y_x$, $\tau y+y_x$ is held fixed, then with increasing λ the zeros of the other move closer to the fixed zero or further away according as* $\int \frac{\partial G}{\partial\lambda} y^2 dx$ *is positive or negative, the integration extending over the interval between the zeros.*

From the preceding theorems we can deduce still further results.

THEOREM XVII. *If* $\int_0^1 \frac{\partial G}{\partial\lambda} y^2 dx$ *has one sign, then between any two consecutive λ's corresponding to solutions of* (41) (43a) *or of* (41) (43d) *there will be one each corresponding to solutions of* (41) (43b) *and* (41) (43c), *and between two consecutive λ's for* (41) (43b) *or* (41) (43c) *there will be one each for* (41) (43a) *and* (41) (43d).

COROLLARY: *Denoting by λ_a and λ_d two adjacent characteristic values for solutions of the problem* (41) (43a), (41) (43d) *respectively, and by λ_b, λ_c and λ_b', λ_c' the next greater and next smaller sets for solutions of the problem* (41) (43b), (41) (43c) *respectively, then if the number of zeros of the solution*

corresponding to λ_a is m, the number corresponding to λ_d is $m-1$. When the integral is positive the number of zeros corresponding to λ_b and λ_c is m, and to λ'_b and λ'_c is $m-1$, while if it is negative, the reverse is true.

If we consider the special form* of (41) where $\dfrac{\partial G}{\partial \lambda} \geqq 0$ and Hypothesis A of § 2 is satisfied while $G(x, -\infty) < 0$ and $= -\infty$ in at least some portion of the interval, we can trace the various values of λ corresponding to successive suites of σ and τ, and state

Theorem XVIII. *There exists under these hypotheses an infinite set of characteristic numbers $\lambda_1 < \lambda_2 < \lambda_3 \ldots$, with limit point at $+\infty$ only, for which exist solutions of (41) (42), the solution $y_n(x)$ corresponding to $\lambda_n (n=1, 2, 3, \ldots)$ having $n-1$ zeros within the interval.*

The results for the orthogonal case (§ 1, II) of the special equation

$$(py_x)_x + (q+\lambda k)y = 0, \tag{46}$$

are contained in the foregoing theorem. In the other cases it is preferable to investigate the zeros through a discussion of those of the equation obtained by the transformation used in proving *Theorem* XVI. We have

$$(\bar{p}\bar{y}_x)_x + (\bar{q}+\lambda\bar{k})\bar{y} = 0, \quad \bar{p}(x) > 0, \quad \bar{y}(0)\bar{y}_x(0) = \bar{y}(1)\bar{y}_x(1) = 0,$$

$$\int_0^1 (\bar{p}\bar{y}_x^2 - \bar{q}\bar{y}^2)\,dx = \int_0^1 \bar{k}\bar{y}^2 dx = \int_0^1 ky^2 dx.$$

If the new equation is of the polar form we see at once that there are precisely two solutions of (46) (42) with n zeros in the interval. If the equation is of the non-definite form there are two integers n_1, n_2, such that for $n < n_1$ there are no solutions of (46) (42), for $n > n_1$ there are at least two, while for $n > n_2$ there are precisely two (Theorems VII, VII A).

§ 7. *Oscillation Theorems for Case II of the Boundary Conditions.*

We shall next consider the exceptional case $B_{42} \equiv \alpha_4\beta_2 - \alpha_2\beta_4 = 0$ for which, as has been seen in § 5, the boundary conditions may be written

$$y(0) = hy(1), \quad hp(0)y_x(0) = p(1)y_x(1), \quad h = \text{constant} \neq 0. \tag{47}$$

Defining the particular solutions $\bar{y}(x, \lambda)$, $\bar{\bar{y}}(x, \lambda)$ by (44), and substituting the solution $y = c_1\bar{y} + c_2\bar{\bar{y}}$ in (47), we obtain equations for c_1 and c_2,

$$c_2 = h[c_1\bar{y}(1, \lambda) + c_2\bar{\bar{y}}(1, \lambda)], \quad hp(0)c_1 = p(1)[c_1\bar{y}_x(1, \lambda) + c_2\bar{\bar{y}}_x(1, \lambda)]. \tag{48}$$

A necessary condition for a solution is the equation

$$D \equiv \begin{vmatrix} h\bar{y}(1, \lambda) & h\bar{\bar{y}}(1, \lambda) - 1 \\ p(1)\bar{y}_x(1, \lambda) - hp(0) & p(1)\bar{\bar{y}}_x(1, \lambda) \end{vmatrix} = 0. \tag{49}$$

* This is a slightly more general condition than that imposed by Birkhoff. *Loc. cit.*

Since for the two solutions we have the well-known formula

$$p(x)[\bar{\bar{y}}_x(x,\lambda)\bar{y}(x,\lambda)-\bar{y}_x(x,\lambda)\bar{\bar{y}}(x,\lambda)]=\text{constant}=-p(0), \tag{50}$$

this condition reduces to

$$h^2p(0)\bar{\bar{y}}(1,\lambda)-2hp(0)+p(1)\bar{y}_x(1,\lambda)=0,$$

from which we get by again using (50)

$$h=\frac{1}{\bar{\bar{y}}(1,\lambda)}\pm\frac{\sqrt{p(1)}\sqrt{-\bar{\bar{y}}_x(1,\lambda)\bar{y}(1,\lambda)}}{\sqrt{p(0)}\bar{\bar{y}}(1,\lambda)}. \tag{51}$$

So long as $\bar{\bar{y}}_x(1,\lambda)$ and $\bar{y}(1,\lambda)$ have opposite signs there will be two values of h for each value of λ; when they have the same sign there will be none. We may distinguish three critical cases for which the solutions pass from real to complex:

(α) If $\bar{y}(1,\lambda)=0$, $\bar{\bar{y}}_x(1,\lambda)\neq 0$, the equation (49) can by means of (50) be written $[h\bar{\bar{y}}(1,\lambda)-1]^2p(1)\bar{\bar{y}}_x(1,\lambda)=0$, and since $\bar{\bar{y}}$ and $\bar{\bar{y}}_x$ can not vanish together, it follows that $h\bar{\bar{y}}(1,\lambda)-1=0$. Hence every element of the determinant D vanishes except that in the lower right-hand corner; it follows then that $c_2=0$. The solution $y(x)\equiv c_1\bar{y}(1,\lambda)$ vanishes at $x=0$, and since $h\neq 0$, we have from (47) that it vanishes also at $x=1$. The function $y(x)$ is therefore a solution of (41) (43a).

(β) If $\bar{y}(1,\lambda)\neq 0$, $\bar{\bar{y}}_x(1,\lambda)=0$, we have in the same way $c_1=0$ and the solution $y\equiv c_2\bar{\bar{y}}$ also a solution of (41) (43d).

(γ) If $\bar{y}(1,\lambda)=0$, $\bar{\bar{y}}_x(1,\lambda)=0$ it follows immediately by the same reasoning that all elements of D vanish. Hence c_1 and c_2 may take on any values and λ is a double characteristic number.

On multiplying equation (41) by y and integrating under the boundary conditions (47) we obtain the relation

$$\int_0^1 py_x^2dx=\int_0^1 Gy^2dx. \tag{52}$$

By substitution in (13) from formulae (47) and the further formulae,

$$h\frac{\partial y(1)}{\partial\lambda}=\frac{\partial y(0)}{\partial\lambda}-\frac{dh}{d\lambda}y(1),\quad \frac{dh}{d\lambda}p(0)y_x(0)+hp(0)\frac{\partial y_x(0)}{\partial\lambda}=p(1)\frac{\partial y_x(1)}{\partial\lambda}, \tag{53}$$

obtained from (47) by differentiation with regard to λ, we obtain the fundamental formula

$$\frac{dh}{d\lambda}=-\frac{h\int_0^1\frac{\partial G}{\partial\lambda}y^2dx}{2p(0)y(0)y_x(0)}=-\frac{h^2\int_0^1\frac{\partial G}{\partial\lambda}y^2dx}{2p(1)y(0)y_x(1)}=-\frac{\int_0^1\frac{\partial G}{\partial\lambda}y^2dx}{2p(0)y(1)y_x(0)}. \tag{54}$$

Let us denote by λ_b a characteristic number of our problem for $h=0$; this is a solution of the sturmian case (41) (43b). Near it, as we know from *Theorem* XVII and its corollary, there is another characteristic number λ_c corresponding to $h=\infty$ and to (41) (43*c*), the two solutions for λ_b and λ_c having the same number of zeros. The parameter value λ_b may be equal to, greater than, or less than λ_c. Of the aggregate of characteristic numbers for the two problems (41) (43a), (41) (43d) let us denote by λ_{ad} the greatest of those smaller than λ_b (or λ_c) and by λ'_{ad} the smallest of those greater. Within the interval λ_{ad}, λ'_{ad} it follows from *Theorem* XVII that there is but one zero and one infinity of h; in other words, but one solution for each of the problems (41) (43b), (41) (43c). On the hypothesis that $\int_0^1 \frac{\partial G}{\partial \lambda} y^2 dx$ has one sign we are now in a position to prove that *in this interval of* λ *the function* $h(\lambda)$ *defined by* (51) *is monotone on both of its branches* (the one including $h=0$ and the other $h=\infty$).

To prove this let us in the first place note that for $h=0$ or $h=\infty$ we have a sturmian case for which existence theorems have already been established, and then show that as h passes through either of these values, $\frac{dh}{d\lambda}$ does not change sign. We see at once from (47) that $y_x(1, \lambda(0))=0$; hence $y(1, \lambda(0)) \neq 0$, and it follows from (47) that $y(0, \lambda(h))$ changes sign with h. A reference to the first part of (54) shows that $\frac{dh}{d\lambda}$ will then retain its sign. In the same way $y(1, \lambda(\infty))=0$, $y_x(1, \lambda(\infty)) \neq 0$; $y_x(0, \lambda(h))$ changes sign as h goes through infinity and $\frac{dh}{d\lambda}$ retains its sign. This establishes the result since it follows from (54) that while $\int_0^1 \frac{\partial G}{\partial \lambda} y^2 dx$ retains one sign the only possibility of $\frac{dh}{d\lambda}$ changing sign is when h or $y(0)$ or $y_x(0)$ changes sign.

By definition $y(0, \lambda(0))=0$, $y_x(1, \lambda(0))=0$, and without loss of generality it may be assumed that $y_x(0, \lambda(0))>0$. The discussion then divides itself into two parts according to

Hypothesis I. $y(1, \lambda(0))>0$. Hypothesis II. $y(1, \lambda(0))<0$.

Under the first hypothesis the number of zeros of y (including that at $x=0$) is even, while under the other it is odd. Roughly speaking, we shall see that a gain or loss of a zero comes when h goes through 0 or ∞. Concerning the function $\lambda(h)$ there is now sufficient data to sketch the graph. *We shall first discuss the problem for the assumption* $\int_0^1 \frac{\partial G}{\partial \lambda} y^2 dx > 0$ *and Hypothesis I.*

It has been proved above that $\frac{dh}{d\lambda}$ can not change sign on either branch of the function. On the branch containing $h=0$, $\lambda=\lambda_b$ we see from (54) that $\frac{dh}{d\lambda}<0$, since for that particular point $y(1)>0$, $y_x(0)>0$. For $h=\infty$, $\lambda=\lambda_c$, the number of zeros is the same as that for $h=0$ (*Corollary, Theorem* XVII) and hence under *Hypothesis* I, $y_x(1)$ and $y(0)$ have opposite signs. From the second part of (54) we see then that $\frac{dh}{d\lambda}>0$ on this other branch. Since at the ends of the interval h is not 0, we know that * in cases (α) and (β) either $y(0)$ or $y_x(0)$ will vanish according as λ_{ad}, λ'_{ad} belong to the problem (41) (43a) or (41) (43d); it follows then from (54) that $\frac{dh}{d\lambda}=\infty$.

Under *Hypothesis* II it is readily shown by the same processes that in an interval which we shall call $\bar{\lambda}_{ad}$, $\bar{\lambda}'_{ad}$ to distinguish it from the other, the branch of the function $\lambda(h)$, which contains $h=0$, is monotone increasing and the branch containing $h=\infty$ is monotone decreasing. As λ increases this form of curve will always alternate with that obtained under *Hypothesis* I. When $\int_0^1 \frac{\partial G}{\partial \lambda} y^2 dx<0$, the second form of curve occurs under *Hypothesis* I, and the first under *Hypothesis* II.

Since by definition the intervals can not overlap, the curves can not overlap. If the upper bound of one interval is the lower bound of the next, we have a double point for λ [cf. (γ) above], and one branch of one curve unites with one of the other to form a function monotone throughout. This is what takes place, for example, in the case of the solutions of the equation † $y_{xx}+\lambda y=0$.

---

* In case (γ), $y(0)$ and $y_x(0)$ may be chosen arbitrarily; cf. next succeeding foot-note.

† For this special equation $\bar{y}\equiv\frac{1}{\sqrt{\lambda}}\sin\sqrt{\lambda}x$, $\bar{\bar{y}}\equiv\cos\sqrt{\lambda}x$, and equation (49) becomes

$$2h-(h^2+1)\cos\sqrt{\lambda}=0,$$

from which we obtain the formula $h=\sec\sqrt{\lambda}\pm\tan\sqrt{\lambda}$. For every positive λ the function h is double-valued. The interval $\lambda_{ad}, \lambda'_{ad}$ is $(2m)^2\pi^2$, $(2m+1)^2\pi^2$, while at $(2m+1/2)^2\pi^2$, h becomes zero on one branch and infinite on the other. The monotone decreasing branch $h_1\equiv\sec\sqrt{\lambda}-\tan\sqrt{\lambda}$ may be considered as joined at both ends to monotone decreasing branches in the next intervals. These intervals $(2m-1)^2\pi^2$, $(2m)^2\pi^2$ and $(2m+1)^2\pi^2$, $(2m+2)^2\pi^2$ are of the type $\bar{\lambda}_{ad}, \bar{\lambda}'_{ad}$. The function h_1 decreases from $+\infty$ to $-\infty$ in the interval $(2m-1/2)^2\pi^2$, $(2m+3/2)^2\pi^2$. The function $h_2\equiv\sec\sqrt{\lambda}+\tan\sqrt{\lambda}$ increases monotonely from $-\infty$ to $+\infty$ in the interval $(m+1/2)^2\pi^2$, $(m+5/2)^2\pi^2$, the branches of adjacent intervals uniting as in the other case. Each of the curves h_1 and h_2 cuts two of the other set orthogonally, the points of intersection occurring at the end points $m^2\pi^2$ of the interval λ_{ad}, λ'_{ad}.

It is now easy to write down oscillation theorems for this case. We note in the first place that under *Hypothesis* I the number of zeros is increased by unity as h goes through zero, and under *Hypothesis* II it is decreased by unity so that for $h>0$ the number of zeros is always even, and for $h<0$ it is odd. There are then two solutions with an even number of zeros in the one case and two with an odd number in the other.

Theorem XIX. *If in an interval of* λ *there exist two integers* m_1, m_2 *positive or zero such that there are solutions of*

$$(py_x)_x+G(x,\lambda)y=0 \tag{55}$$

and (43*a*) *with* m_1 *and* m_2 *zeros respectively within the interval* 0, 1, *then under the hypothesis that* $\int_0^1 \frac{\partial G}{\partial\lambda} y^2 dx$ *has one sign throughout, there are for* $h>0$ *two solutions of* (55) (47) *when* m *is even, and none when* m *is odd* $(m_1 \leq m \leq m_2)$: *for* $h<0$ *there are two solutions when* m *is odd aud none when* m *is even.*

The special case where $\frac{\partial G}{\partial\lambda}>0$ and G becomes negative for $\lambda=-\infty$, and moreover negatively infinite in at least a part of the interval and G becomes positively infinite for $\lambda=+\infty$ in at least a part of the interval, includes the case discussed by Birkhoff and the detailed theorems derived by him hold also here.*

The orthogonal problem (§ 1, II) of the equation

$$(py_x)_x+(q+\lambda k)y=0 \tag{56}$$

is contained in the special case just discussed. In discussing the polar case we note that it follows from the special case of the formula (52),

$$0<\int_0^1(py_x^2-qy^2)\,dx=\lambda\int_0^1 ky^2dx, \tag{57}$$

that $\int_0^1 \frac{\partial G}{\partial\lambda} y^2 dx \equiv \int_0^1 ky^2dx$ has the same sign as λ. Hence $\lambda=0$ is not included in the range of values (unless one includes the solution $y\equiv0$).

Theorem XX. *In the polar case of equation* (56) *there are precisely two solutions satisfying the boundary conditions* (47) *and oscillating* n *times* $(n=0, 1, 2, \ldots)$.

In the non-definite case we have from *Theorem* VII that for sufficiently large values of $|\lambda|$ the integral on the left of (57) is positive and hence we can state

---

* Case II corresponds to III, p. 269 in Birkhoff's article, *loc. cit.*

THEOREM XXI. *For the non-definite case of equation* (56) *there exist two integers* n_1, n_2 $(n_2 \geqq n_1 \geqq 0)$ *such that for* $n < n_1$ *there is no solution of* (56) (47) *with* n *zeros in the interval; for* $n > n_1$ *there are at least two, while for* $n \geqq n_2$ *there are precisely two.*

§ 8. *Oscillation Theorems for Case III of the Boundary Conditions.*

There remains the most general of the three normal forms obtained for the boundary conditions, viz.:

$$y(0) = lp(1) y_x(1), \quad lp(0) y_x(0) = -y(1), \quad l = \text{constant} \neq 0. \tag{58}$$

Since the discussion follows the same lines as that of the preceding section it will be abbreviated. If we define the two particular solutions $\bar{y}(x, \lambda)$, $\bar{\bar{y}}(x, \lambda)$ as in (44), we get in place of formulae (49), (51),

$$\left. \begin{aligned} D' &\equiv \begin{vmatrix} lp(1)\bar{y}_x(1,\lambda) & lp(1)\bar{\bar{y}}_x(1,\lambda)-1 \\ \bar{y}(1,\lambda)+lp(0) & \bar{\bar{y}}(1,\lambda) \end{vmatrix} = 0, \\ l &= \frac{1}{p(1)\bar{\bar{y}}_x(1,\lambda)} \pm \frac{\sqrt{\bar{y}_x(1,\lambda)\bar{\bar{y}}(1,\lambda)}}{\sqrt{p(0)\,p(1)\bar{\bar{y}}_x(1,\lambda)}}\,. \end{aligned} \right\} \tag{59}$$

The critical values are when $\bar{y}_x(1, \lambda) = 0$ and $\bar{\bar{y}}(1, \lambda) = 0$ and the cases may be classified as before.

(α) If $\bar{\bar{y}}(1, \lambda) = 0$, $\bar{y}_x(1, \lambda) \neq 0$, equation (59) can be written

$$\bar{y}(1,\lambda)\,[lp(1)\bar{\bar{y}}_x(1,\lambda)-1]^2 = 0 \quad \text{or} \quad lp(1)\bar{\bar{y}}_x(1,\lambda) = 1.$$

It follows that $c_1 = 0$ and the solution $y(x, \lambda) \equiv c_2 \bar{\bar{y}}(x, \lambda)$ is a solution of the equation with the sturmian boundary condition (43c).

(β) If $\bar{\bar{y}}(1, \lambda) \neq 0$, $\bar{y}_x(1, \lambda) = 0$ we have $c_2 = 0$ and $y \equiv c_1 \bar{y}$, which is a solution of (41) (43b).

(γ) If $\bar{\bar{y}}(1, \lambda) = 0$, $\bar{y}_x(1, \lambda) = 0$, it follows as before that D' vanishes identically and λ is a double parameter value.

To replace formula (53) and (54) we have

$$\frac{\partial y(0)}{\partial \lambda} = \frac{dl}{d\lambda} p(1) y_x(1) + lp(1) \frac{\partial y_x(1)}{\partial \lambda},$$

$$lp(0) \frac{\partial y_x(0)}{\partial \lambda} + \frac{dl}{d\lambda} p(0) y_x(0) = -\frac{\partial y(1)}{\partial \lambda},$$

$$\frac{dl}{d\lambda} = -\frac{l \int_0^1 \frac{\partial G}{\partial \lambda} y^2 dx}{2p(0) y_x(0) y(0)} = \frac{l^2 \int_0^1 \frac{\partial G}{\partial \lambda} y^2 dx}{2y(1) y(0)} = \frac{-\int_0^1 \frac{\partial G}{\partial \lambda} y^2 dx}{2p(0) y_x(0) p(1) y_x(1)}\,. \tag{60}$$

It is readily shown as in § 7 that $\frac{dl}{d\lambda}$ does not change sign as l goes through the values 0 or ∞. Let us denote by λ_a, λ_d two adjacent characteristic numbers for the cases $l=0$ and $l=\infty$, respectively (in other words for (41) (43a), (41) (43d)), by λ_{bc} the greatest of the aggregate of characteristic numbers for (41) (43b) and (41) (43c) which are smaller than λ_a, and by λ'_{bc} the smallest of the aggregate larger. It is readily shown as in the previous section that within the interval λ_{bc}, λ'_{bc}, l is a monotone function on the branch containing $l=0$ and on the branch containing $h=\infty$. Let us consider first the case $l=0$; then $y(0, \lambda(0))=0$, $y(1, \lambda(0))=0$ and we can assume $y_x(0, \lambda(0))>0$. There will be two cases to distinguish according as we make

Hypothesis I. $y_x(1, \lambda(0))>0$; Hypothesis II. $y_x(1, \lambda(0))<0$.

Under *Hypothesis* I the number of zeros is always odd, and when l goes through zero from negative to positive it may be seen from (58) that two zeros of the solution $y(x)$ are lost. Under *Hypothesis* II the number of zeros is even, and, as l increases through zero, two zeros of $y(x)$ are gained. It follows from the last part of (60) that the integral $\int_0^1 \frac{\partial G}{\partial \lambda} y^2 dx$ and $\frac{dl(0)}{d\lambda}$ have opposite signs. This fixes the sign of $\frac{dl}{d\lambda}$ on one branch. To fix the sign of $\frac{dl}{d\lambda}$ on the other branch let us consider $\lambda(\infty)\equiv\lambda_d$. Under *Hypothesis* I we can argue from the corollary to *Theorem* XVII that $y(0, \lambda(\infty))$ and $y(1, \lambda(\infty))$ have the same sign, and from the third part of (60) that when $\int_0^1 \frac{\partial G}{\partial \lambda} y^2 dx>0$, $\frac{dl(\infty)}{d\lambda}$ is positive. Under *Hypothesis* II we can easily prove that the situation is reversed, $\frac{dl}{d\lambda}$ being positive on the branch through $h=0$, and negative on the other. When $\int_0^1 \frac{\partial G}{\partial \lambda} y^2 dx<0$ the two types of curves are interchanged. With increasing λ we see then that in both cases curves of much the same form as the two varieties in the preceding section alternate with one another. If adjacent intervals have the same end-point, a branch from the one will unite with a branch of the other.*

From these data various theorems may be deduced, of which the following is typical.

* For the special equation $y_{xx}+\lambda y=0$ one can set up the formula for $l(\lambda)$ as in foot-note, p. 312. It is found that $l=-\csc\sqrt{\lambda}\pm\cot\sqrt{\lambda}/\sqrt{\lambda}$ and that these monotone functions extend from $-\infty$ to $+\infty$, each of the one set cutting two of the other set in double points $\lambda=(n+1/2)^2\pi^2$ of $l(\lambda)$.

THEOREM XXII. *If in an interval of* λ *there exist two integers* $\mu_1 < \mu_2$, *positive or zero, such that there are solutions of* (55) (43*a*) *with* μ_1 *and* μ_2 *zeros respectively within the interval* 0, 1, *then*

(1) *Under the hypothesis that* $\int_0^1 \frac{\partial G}{\partial \lambda} y^2 dx$ *be positive, there are, when* l *is positive, two solutions of* (55) (58) *which have* $2m$ *or* $2m-1$ *zeros in the interval* $(\mu_1 < 2m-1 < 2m < \mu_2)$ *and there are, when* l *is negative, two solutions which have* $2m$ *or* $2m+1$ *zeros* $(\mu_1 < 2m < 2m+1 < \mu_2)$.

(2) *Under the hypothesis that* $\int_0^1 \frac{\partial G}{\partial \lambda} y^2 dx$ *be negative, there are, when* l *is positive, two solutions of* (55) (58) *which have* $2m$ *or* $2m+1$ *zeros in the interval* $(\mu_1 < 2m < 2m+1 < \mu_2)$ *and there are, when* l *is negative, two solutions which have* $2m$ *or* $2m-1$ *zeros* $(\mu_1 < 2m-1 < 2m < \mu_2)$.

One can make the results of this theorem more specific by giving the conditions necessary to characterize the branch of the function $l(\lambda)$ which is involved. Birkhoff has done this for the special case treated by him (*loc. cit.*, p. 269, I, II), but we shall content ourselves with stating that the same classification may be made in the general case treated here.

Theorems for the special equation (56) analogous to *Theorems* XX, XXI, can be at once written down.

BROWN UNIVERSITY, *January*, 1918.